

GenCore version 5.1.6
Copyright (c) 1993-2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:20:35 ; search time 73 seconds
(without alignments)
1122.591 Million cell updates/sec

Title: US-09-898-165B-7

Perfect score: 615

Sequence: 1 MSGIKKQKTENQOKSTNVVY.....MAPKAMKVITDYVRSEMDKN 615

Scoring table: OLIGO

Gapop 60.0 , capext 60.0

Searched: 908470 seqs, 133250630 residues

Word size : 10

Total number of hits satisfying chosen parameters: 61

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries.

Database :

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	288	46.8	648	22 AAU32438	Novel human secret
2	221	35.9	619	21 AAU79214	Human transferase
3	65	10.6	65	22 AB40266	Peptide #7772 enco
4	65	10.6	65	22 AB24683	Protein #6682 enco
5	65	10.6	65	22 AB24683	Human brain expres
6	65	10.6	65	22 AAU73760	Human bone marrow
7	65	10.6	65	22 AAU73760	Peptide #7984 enco
8	65	10.6	65	23 ABG43644	Human peptide enco
9	55	8.9	133	23 ABP41780	Human ovarian anti
10	38	6.2	624	19 AAU70494	Human disease rela

11	38	6.2	624	20	AAU22349	Human APS kinase/A
12	38	6.2	625	20	AAU67882	Human secreted pro
13	36	5.9	36	19	AAU75032	Fragment of human
14	22	3.6	373	21	AAG44041	Arabidopsis thalia
15	22	3.6	460	21	AAG44040	Arabidopsis thalia
16	22	3.6	463	21	AAG44039	Arabidopsis thalia
17	21	3.4	373	21	AAG21123	Arabidopsis thalia
18	21	3.4	385	21	AAG21122	Arabidopsis thalia
19	21	3.4	396	21	AAG43668	Arabidopsis thalia
20	21	3.4	408	21	AAG43667	Arabidopsis thalia
21	21	3.4	425	21	AAG09596	Arabidopsis thalia
22	21	3.4	446	21	AAG21121	Arabidopsis thalia
23	21	3.4	462	21	AAG30254	Arabidopsis thalia
24	21	3.4	465	21	AAG30253	Arabidopsis thalia
25	21	3.4	469	21	AAG43666	Arabidopsis thalia
26	21	3.4	473	21	AAG09595	Arabidopsis thalia
27	21	3.4	476	21	AAG09594	Arabidopsis thalia
28	21	3.4	510	21	AAG30252	Arabidopsis thalia
29	20	3.3	635	22	ABU71650	Drosophila melanog
30	15	2.4	140	21	AAG20076	Arabidopsis thalia
31	15	2.4	166	21	AAG20075	Arabidopsis thalia
32	15	2.4	174	21	AAG20074	Arabidopsis thalia
33	14	2.3	224	21	AAU44788	Corn Adenylsulph
34	12	2.0	139	21	AAG48931	Arabidopsis thalia
35	12	2.0	151	21	AAG48930	Arabidopsis thalia
36	12	2.0	164	21	AAG48929	Arabidopsis thalia
37	12	2.0	236	21	AAG14478	Arabidopsis thalia
38	12	2.0	251	21	AAG11017	Arabidopsis thalia
39	12	2.0	251	21	AAG38348	Arabidopsis thalia
40	12	2.0	259	21	AAG14477	Arabidopsis thalia
41	12	2.0	263	21	AAG11016	Arabidopsis thalia
42	12	2.0	263	21	AAG38347	Arabidopsis thalia
43	12	2.0	272	21	AAG14476	Arabidopsis thalia
44	12	2.0	276	21	AAG11015	Arabidopsis thalia
45	12	2.0	276	21	AAG38346	Arabidopsis thalia
46	12	2.0	293	21	AAG29425	Arabidopsis thalia
47	12	2.0	293	21	AAU77957	A. thaliana enviro
48	11	1.8	161	21	AAU53456	Arabidopsis thalia
49	11	1.8	204	23	ABP39672	Staphylococcus epi
50	11	1.8	208	21	AAG35788	Arabidopsis thalia
51	11	1.8	237	21	AAG53120	Arabidopsis thalia
52	11	1.8	246	21	AAU44792	Wheat Adenylsulph
53	11	1.8	252	21	AAG14580	Arabidopsis thalia
54	11	1.8	252	21	AAG47458	Arabidopsis thalia
55	11	1.8	305	21	AAG14579	Arabidopsis thalia
56	11	1.8	305	21	AAG47457	Arabidopsis thalia
57	11	1.8	310	21	AAG14578	Arabidopsis thalia
58	11	1.8	310	21	AAG47456	Arabidopsis thalia
59	10	1.6	201	22	AAG58409	Escherichia coli p
60	10	1.6	311	21	AAU44791	Soybean Adenylsul
61	10	1.6	343	21	AAU44789	Corn Adenylsulph

ALIGNMENTS

RESULT 1

AAU32438
ID AAU32438 standard; Protein: 648 AA.
XX XX
AC AAU32438;
XX XX

DT 18-DEC-2001 (first entry)

XX XX Novel human secreted protein #2929.

Human: vaccination; gene therapy; nutritional supplement;
stem cell proliferation; hematopoiesis; nerve tissue regeneration;
immune suppression; immune stimulation; anti-inflammatory; leukaemia.
Homo sapiens.
WO200179449-A2.

XX PD 25-OCT-2001.
 XX XX
 XX PF 16-APR-2001; 2001WO-US08656.
 XX XX
 XX PR 18-APR-2000; 2000US-0552929.
 XX PR 26-JAN-2001; 2001US-0770160.
 XX (HYSE-) HYSEQ INC.
 XX PA
 XX XX
 XX PI Tang YT, Liu C, Drmanac RT;
 XX PI WPI; 2001-611725/70.
 XX DR
 XX XX
 XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 XX PT vaccination, testing and therapy
 XX XX
 XX PS Claim 20; Page 613; 765pp; English.
 XX XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising
 CC the nucleic acids encoding the polypeptides and cells genetically
 CC engineered to express them are also useful for producing the proteins.
 CC The proteins are useful in genetic vaccination, testing and
 CC therapy, and can be used as nutritional supplements. They may be used to
 CC increase stem cell proliferation; to regulate haematopoiesis; and in
 CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
 CC immune suppression and/or stimulation; as anti-inflammatory agents; and
 CC in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid
 CC sequences of novel human secreted proteins of the invention.
 XX XX
 SQ Sequence 648 AA;
 Query Match 46.8%; Score 288; DB 22; Length 648;
 Best Local Similarity 100.0%; Pred. No. 7.8e-282;
 Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSGIKKQKTEHQKSTNNVYQAHVSRNRKRGQVVGTRGGFGCTVWLTLGLSGAGKTTISF 60
 DB 27 MSGIKKQKTEHQKSTNNVYQAHVSRNRKRGQVVGTRGGFGCTVWLTLGLSGAGKTTISF 86
 QY 61 ALEEYLVSHPAIPCYSLDGDVNRHGLNRNLGFSFGDREENIRIRAEVAKLFADAGLCVITS 120
 DB 87 ALEEYLVSHPAIPCYSLDGDVNRHGLNRNLGFSFGDREENIRIRAEVAKLFADAGLCVITS 146
 QY 121 FISPFADRENARKIHESAGLPPEFIFVDAPLNICESRDVKGLYKARAGEIKGFTGIDS 180
 DB 147 FISPFADRENARKIHESAGLPPEFIFVDAPLNICESRDVKGLYKARAGEIKGFTGIDS 206
 QY 181 DYEKPTPEPVRLKTNLSVSDCVHQVVELLQEQNIIVPYTIKDIHELFPENKLDHVAE 240
 DB 207 DYEKPTPEPVRLKTNLSVSDCVHQVVELLQEQNIIVPYTIKDIHELFPENKLDHVAE 266
 QY 241 AETPLSLSTIKLDLQWVQLSEGWATPLKGFMRKEYLQVMHFDTLDD 288
 DB 267 AETPLSLSTIKLDLQWVQLSEGWATPLKGFMRKEYLQVMHFDTLDD 314

RESULT 2

AA79214
 ID AAY79214 standard; Protein: 619 AA.
 XX
 XX AAY79214;
 XX
 XX DT 19-JUN-2000 (first entry)
 XX XX
 CC Human transferase TRNSFS-6.

XX Transferase; TRNSFS-6; human; antitumour; cancer;
 KW gastrointestinal disorder; developmental disorder;
 KW genetic disorder; neurological disorder; reproductive disorder;
 KW smooth muscle disorder; immunological disorder; inflammation;
 KW diagnosis; therapy: ATP sulfurylase/APS kinase 2.
 XX

OS Homo sapiens.

Key	Location/Qualifiers
Modified-site 92	/note= "potential O-phosphorylation"
Modified-site 176	/note= "potential O-phosphorylation"
Modified-site 180	/note= "potential O-phosphorylation"
Modified-site 198	/note= "potential O-phosphorylation"
Modified-site 250	/note= "potential O-phosphorylation"
Modified-site 285	/note= "potential O-phosphorylation"
Modified-site 308	/note= "potential O-phosphorylation"
Modified-site 313	/note= "potential O-phosphorylation"
Modified-site 394	/note= "potential O-phosphorylation"
Modified-site 453	/note= "potential O-phosphorylation"
Modified-site 527	/note= "potential O-phosphorylation"
Modified-site 548	/note= "potential O-phosphorylation"
Modified-site 430	/note= "potential O-phosphorylation"
Modified-site 439	/note= "potential O-phosphorylation"
Modified-site 195	/note= "potential N-glycosylation"
Modified-site 298	/note= "potential N-glycosylation"
Binding-site 49..56	/note= "ATP/GTP binding site (P-loop)"

WO200014251-A2.

16-MAR-2000.

09-SEP-1999; 99WO-US20989.

10-SEP-1998; 98US-0150657.

04-NOV-1998; 98US-0186779.

11-MAY-1999; 99US-0133642.

(INCY-) INCYTE PHARM INC.

Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;

Hillman JL, Azimzai Y;

WPI; 2000-256996/22.

N-PSDB; AA294206.

Human transferase proteins useful for preventing, diagnosing and
 treating cancers and developmental, gastrointestinal, genetic,
 immunological, neurological, reproductive and smooth muscle disorders -

Claim 1; Page 84-85; 113pp; English.

The present sequence is that of human transferase TRNSFS-6, 1 of
 15 claimed novel human transferase proteins of the invention (see
 AA79209-23). The sequence was deduced from a cDNA clone (see
 AA294206) isolated from a foetal kidney library. It shows homology

CC to human ATP sulfurylase/APS kinase 2. TRNSFS-6 is expressed in
 CC cardiovascular, gastrointestinal, reproductive and endocrine
 CC tissues, especially those associated with cancer, inflammation and
 CC cell proliferation. The new human transferases and polynucleotides
 CC can be used in the diagnosis, prevention and treatment of cancer,
 CC developmental disorders, gastrointestinal disorders, genetic
 CC disorders, immunological disorders, neurological disorders,
 CC reproductive disorders, and smooth muscle disorders. The
 CC polypeptides can also be used to raise antibodies, and to screen
 CC for agonists and antagonists of transferase activity.

XX Sequence 619 AA;
 Query Match 35.9%; Score: 221; DB 21; Length 619;
 Best Local Similarity 99.7%; Pred. No: 3.7e-214;
 Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 289 DGVINNSIPVLVPSAEDKTRLECCSFVLAHGRVAILRDAAEFYEHKKEECSRWGT 348
 DB 294 DGVINNSIPVLVPSAEDKTRLECCSFVLAHGRVAILRDAAEFYEHKKEECSRWGT 353
 QY 349 TCTKHPHKIMVMSGDLVGGDLQVLEKIRWNDGLDQYRLTPLELKOKCKEMNADAVFAF 408
 DB 354 TCTKHPHKIMVMSGDLVGGDLQVLEKIRWNDGLDQYRLTPLELKOKCKEMNADAVFAF 413
 QY 409 QLRNPVNHGHALLMQDTCRRLLERGYKHPVLLHPLGGWTKDDVPLDMRMKHAVALLEE 468
 DB 414 QLRNPVNHGHALLMQDTCRRLLERGYKHPVLLHPLGGWTKDDVPLDMRMKHAVALLEE 473
 QY 469 GVLDPKSTIVAIFPSPMLYAGPTEVQWHCHSRMIAGANFYIVGRDPAGMPHPTKKDLYE 528
 DB 474 GVLDPKSTIVAIFPSPMLYAGPTEVQWHCHSRMIAGANFYIVGRDPAGMPHPTKKDLYE 533
 QY 529 PTHGGKVLSPAGLTSVEIIPFVAAYNKAKKAMDFYDPAHNEFFDISGTRMKKLAREG 588
 DB 534 PTHGGKVLSPAGLTSVEIIPFVAAYNKAKKAMDFYDPAHNEFFDISGTRMKKLAREG 593
 QY 589 ENPPDGFMAPKANKVLTDTYRS 610
 DB 594 ENPPDGFMAPKANKVLTDTYRS 615

RESULT 3
 ABB40266
 ID ABB40266 standard; Peptide; 65 AA.
 XX ABB40266;
 AC ABB40266;
 DT 04-FEB-2002 (first entry)
 DE Peptide #7772 encoded by human foetal liver single exon probe.
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX Homo sapiens.
 XX WO200157277-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00669.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488899/53.

XX WPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human foetal liver.
 PT Claim 27; SEQ ID NO 32901; 639pp + sequence listing; English.
 XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 65 AA;
 Query Match 10.6%; Score 65; DB 22; Length 65;
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EEVLVSHAIPCYSLDGDVNRHGLNRNLGFSFGDRENNIRIAEVAKLFPADAGLVCTTSFI 122
 DB 1 EEVLVSHAIPCYSLDGDVNRHGLNRNLGFSFGDRENNIRIAEVAKLFPADAGLVCTTSFI 122
 QY 123 SPEAK 127
 DB 61 SPEAK 65

RESULT 4
 ABB24683
 ID ABB24683 standard; Protein; 65 AA.
 XX ABB24683;
 AC ABB24683;
 DT 23-JAN-2002 (first entry)
 DE Protein #682 encoded by probe for measuring heart cell gene expression.
 XX Human; gene expression; heart; microarray; vascular system;
 XX cardiovascular disease; hypertension; cardiac arrhythmia;
 XX congenital heart disease.
 XX Homo sapiens.
 XX WO200157274-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00666.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488899/53.

Single exon nucleic acid probes for analyzing gene expression in human
 hearts -
 Claim 15; SEQ ID No 26453; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21535-ABA411305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences

XX Sequence 65 AA;
 SQ

Query Match 10.6%; Score 65; DB 22; Length 65;
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EYLVSHAIPCVSLDGNVHRHGLNRLGSPGDRENIIRIAEVAKLFDAGLVCITSTFI 122
 DB 1 EYLVSHAIPCVSLDGNVHRHGLNRLGSPGDRENIIRIAEVAKLFDAGLVCITSTFI 60
 QY 123 SPFAK 127
 DB 61 SPFAK 65

RESULT 5
 AAM61056
 ID AAM61056 standard; Protein; 65 AA.
 XX AAM61056;
 AC
 XX
 XX 05-NOV-2001 (first entry)
 DT
 XX
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33161.
 XX
 XX Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 PD
 XX 09-AUG-2001.
 PF
 XX 30-JAN-2001; 2001WO-US00667.
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 PT
 XX
 PS Example 4; SEQ ID NO: 33161; 650pp + Sequence Listing; English.
 XX
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.

XX Sequence 65 AA;
 SQ

Query Match 10.6%; Score 65; DB 22; Length 65;
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EYLVSHAIPCVSLDGNVHRHGLNRLGSPGDRENIIRIAEVAKLFDAGLVCITSTFI 122
 DB 1 EYLVSHAIPCVSLDGNVHRHGLNRLGSPGDRENIIRIAEVAKLFDAGLVCITSTFI 60
 QY 123 SPFAK 127
 DB 61 SPFAK 65

RESULT 6
 AAM73760
 ID AAM73760 standard; Protein; 65 AA.
 XX AAM73760;
 AC
 XX
 XX 06-NOV-2001 (first entry)
 DT
 XX
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 34066.
 XX
 XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 PD
 XX 09-AUG-2001.
 PF
 XX 30-JAN-2001; 2001WO-US00668.
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 PT
 XX
 PS Example 4; SEQ ID NO: 34066; 658pp + Sequence Listing; English.
 XX
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.

XX Sequence 65 AA;
 SQ

Query Match 10.6%; Score 65; DB 22; Length 65;
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 EYLVSHAIPCYSLDGDNVRHGLNRNLGFSFGDREENIRRTAEVAKLFADAGLVCIITSFI 122
 Db 1 EYLVSHAIPCYSLDGDNVRHGLNRNLGFSFGDREENIRRTAEVAKLFADAGLVCIITSFI 60
 Qy 123 SPFAK 127
 Db 61 SPFAK 65

RESULT 7
 AAM33947
 ID AAM33947 standard; Protein: 65 AA.
 XX
 AC AAM33947;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Peptide #7984 encoded by probe for measuring placental gene expression.

Probe; microarray; human; placenta; antenatal diagnosis;
 genetic disorder.
 OS Homo sapiens.
 XX WO200157272-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488897/53.
 XX
 PT Human genome-derived single exon-nucleic acid probes useful for
 PT analyzing gene expression in human placenta
 PS
 PS Claim 27; SEQ ID No 34216; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP:
 CC see AA1315-AA157346). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 XX
 XX Sequence 65 AA;
 SQ
 Query Match 10.6%; Score 65; DB 22; Length 65;
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 63 EYLVSHAIPCYSLDGDNVRHGLNRNLGFSFGDREENIRRTAEVAKLFADAGLVCIITSFI 122
 Db 1 EYLVSHAIPCYSLDGDNVRHGLNRNLGFSFGDREENIRRTAEVAKLFADAGLVCIITSFI 60
 Qy 123 SPFAK 127
 Db 61 SPFAK 65

RESULT 8
 ABG43644

ID ABG43644 standard; Peptide: 65 AA.
 XX
 AC ABG43644;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 33309.
 XX
 KW Human: single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberculous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS
 PS Claim 27; SEQ ID No 33309; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease

CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemolysis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension,
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 65 AA;

Query Match 10.6%; Score 65; DB 23; Length 65;
 Best Local Similarity 100.0%; Pred. No. 2.1e-57;
 Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EYLVSHPICSLDGNVRHGLNRNLGSPGDRNIRIAEVAKLFDAGLVCITSI 122
 Db 1 EYLVSHPICSLDGNVRHGLNRNLGSPGDRNIRIAEVAKLFDAGLVCITSI 60

QY 123 SPEAK 127

Db 61 SPEAK 65

RESULT 9

ABP41780

ID ABP41780 standard; Protein; 133 AA.

XX AC ABP41780;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HMEJA45, SEQ ID NO:2912.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive.

XX OS Homo sapiens.

XX PN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-US18569.

XX PR 07-JUN-2000; 2000US-209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-147878/19.

XX DR N-PSDB; ABQ54857.

XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -

XX PS Claim 11; SEQ ID No 2912; 2922pp; English.

XX

CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis), and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 133 AA;

Query Match 8.9%; Score 55; DB 23; Length 133;
 Best Local Similarity 100.0%; Pred. No. 5.1e-47;

Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGIKKQKTENQOKSTNVVYQAHVSRNKRGOVVGTRGGRCGTVMLTGLSGAGK 55

Db 56 MSGIKKQKTENQOKSTNVVYQAHVSRNKRGOVVGTRGGRCGTVMLTGLSGAGK 110

RESULT 10

AAW70494

ID AAW70494 standard; Protein; 624 AA.

XX AC AAW70494;

XX DT 29-DEC-1998 (first entry)

XX DE Human disease related nucleotide kinase-2 (DRNK-2) protein sequence.

XX KW Human disease related nucleotide kinase-2; DRNK-2; deoxyguanosine kinase;
 KW p21ras; cell proliferation; oncogenesis; cancer; PAPS;
 KW immune disorder; neurological dysfunction; gene therapy;
 KW adenosine 3'-phosphate 5'-phosphosulfate-dependent enzyme motif;
 KW ATP pyrophosphatase PP-motif.

XX OS Homo sapiens.

XX FH Key

XX FT Modified-site 27

XX FT Domain /note= "N-glycosylated"

XX FT Region 59..86

XX FT Region /note= "ATP/GTP binding site (p-loop)"

XX FT Region 176..187

XX FT Modified-site /note= "adenosine 3'-phosphate 5'-phosphosulfate

XX FT Modified-site (PAPS)-dependent enzyme motif"

XX FT Region /note= "N-glycosylated"

XX FT Region 411..433

XX FT /note= "ATP pyrophosphatase PP-motif"

XX PN US5817482-A.

XX PD 06-OCT-1998.
XX XX
XX PF 20-JUN-1997; 97US-0879561.
XX PR 20-JUN-1997; 97US-0879561.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Bandman O, Corley NC, Guegler KJ, Hawkins PR, Hillman JL;
XX DR WPI: 1998-556388/47.
XX DR N-PSDB: AAV33482.
XX PT Nucleic acids encoding deoxyguanosine kinase - useful for
XX PT recombinant production of the enzyme for treating diseases caused by
XX PT lack of the enzyme e.g. cancers caused through loss of enzyme
XX PT function
XX XX
XX PS Examples; Fig 2A-2G; 53pp; English.
XX CC The present sequence represents a human disease related nucleotide
XX CC kinase-2 (DRNK-2) protein sequence. The DNKR-2 encoding DNA sequence
XX CC was first identified in Incyte Clone 373887 from the lung tissue cDNA
XX CC library (LUNGNOT 02). The DNKR-2 DNA sequence is useful for the
XX CC production of the corresponding recombinant enzyme. The invention
XX CC provides DRNK enzymes, which are deoxyguanosine kinases, which catalyse
XX CC the transfer of a terminal phosphate from adenosine triphosphate (ATP)
XX CC or guanine triphosphate (GTP) to guanosine or guanidine in the regulation
XX CC of cellular levels of GTP and its corresponding nucleoside triphosphate.
XX CC As GTP levels are known to control the activity of certain oncogenic
XX CC proteins e.g. p21ras, a protein involved in cell proliferation and
XX CC oncogenesis, suppression of the enzyme activity causes high ratios of
XX CC GTP:GDP, promoting oncogenesis. Therefore, diseases (e.g. cancers,
XX CC immune disorders and neurological dysfunction) related to this lack of
XX CC activity may be prevented or treated with the recombinant enzyme, or by
XX CC gene therapy based strategies. Anti-sense constructs of the DNKR
XX CC encoding nucleic acids may also be used for inhibition of over-expression
XX CC of the enzyme.
XX SQ Sequence 624 AA;
XX
XX Query Match 6.2%; Score 38; DB 19; Length 624;
XX Best Local Similarity 100.0%; Pred. No. 3e-29;
XX Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 20 YQAHVHSNRKRGQVVGTRGGRGCTVWLTGLSGAGKTT 57
XX |
XX Db 30 YQAHVHSNRKRGQVVGTRGGRGCTVWLTGLSGAGKTT 67
XX
XX RESULT 11
XX AAY22349
XX ID AAY22349 standard; Protein; 624 AA.
XX XX
XX AC AAY22349;
XX XX
XX DT 24-SEP-1999 (first entry)
XX XX
XX DE Human APS kinase/ATP sulphurylase protein.
XX XX
XX KW APS kinase/ATP sulphurylase; human; PAPS production;
XX KW 3'-phosphoadenosine 5'-phosphosulphate.
XX XX
XX OS Homo sapiens.
XX PN JP11187883-A.
XX XX
XX PD 13-JUL-1999.
XX XX
XX PF 26-DEC-1997; 97JP-0360387.
XX XX
XX PF 26-DEC-1997; 97JP-0360387.

XX PA (NIRA) UNITIKA LTD.
XX PA (HUMA-) ZH HUMAN SCI SHINKO ZAIDAN.
XX XX
XX DR WPI: 1999-451549/38.
XX DR N-PSDB: AAX84897.
XX XX
XX PT New human-derived APS kinase/ATP sulfurylase gene - useful for
XX PT large scale preparation of 3'-phosphoadenosine 5'-phosphosulfate
XX PT (PAPS)
XX XX
XX PS Claim 1: Page 6-7; 9pp; Japanese.
XX XX
XX CC This sequence is the human-derived APS kinase/ATP sulphurylase of
XX CC the invention. The enzyme may be used to prepare 3'-phosphoadenosine
XX CC 5'-phosphosulphate (PAPS) on a large scale.
XX XX
XX SQ Sequence 624 AA;
XX
XX Query Match 6.2%; Score 38; DB 20; Length 624;
XX Best Local Similarity 100.0%; Pred. No. 3e-29;
XX Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 20 YQAHVHSNRKRGQVVGTRGGRGCTVWLTGLSGAGKTT 57
XX |
XX Db 30 YQAHVHSNRKRGQVVGTRGGRGCTVWLTGLSGAGKTT 67
XX
XX RESULT 12
XX AAW67882
XX ID AAW67882 standard; Protein; 625 AA.
XX XX
XX AC AAW67882;
XX XX
XX DT 25-MAR-1999 (first entry)
XX XX
XX DE Human secreted protein encoded by gene 76 clone HOSFD58.
XX XX
XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
XX KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
XX KW developmental abnormality; foetal deficiency; blood; allergy; renal;
XX KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
XX KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
XX KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
XX KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
XX KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 625
XX FT /label= unknown
XX XX
XX PN WO9842738-A1.
XX XX
XX PD 01-OCT-1998.
XX XX
XX PF 19-MAR-1998; 98WO-US05311.
XX XX
XX PR 30-MAY-1997; 97US-0050937.
XX PR 21-MAR-1997; 97US-0041276.
XX PR 21-MAR-1997; 97US-0041277.
XX PR 21-MAR-1997; 97US-0041281.
XX PR 21-MAR-1997; 97US-0042344.
XX PR 30-MAY-1997; 97US-0048069.
XX PR 30-MAY-1997; 97US-0048094.
XX PR 30-MAY-1997; 97US-0048095.
XX PR 30-MAY-1997; 97US-0048096.
XX PR 30-MAY-1997; 97US-0048099.
XX PR 30-MAY-1997; 97US-0048131.
XX PR 30-MAY-1997; 97US-0048135.
XX PR 30-MAY-1997; 97US-0048154.
XX PR 30-MAY-1997; 97US-0048160.

```
PR 30-MAY-1997; 97US-0048186.
PR 30-MAY-1997; 97US-0048187.
PR 30-MAY-1997; 97US-0048188.
PR 30-MAY-1997; 97US-0048350.
PR 30-MAY-1997; 97US-0048351.
PR 30-MAY-1997; 97US-0048352.
PR 30-MAY-1997; 97US-0048353.
PR 30-MAY-1997; 97US-0048355.
PR 05-AUG-1997; 97US-0054804.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brewer LA, Duan R, Ebner R, Ferrie AM, Florence KA;
XX Greene JM, Hu JS, Lafleur DW, Moore PA, NI, J, Olsen HS;
XX Rosen CA, Ruben SM, Shi Y, Young P;
XX
XX WPI: 1999-070066/06.
XX N-PSDB: AAX00686.
XX
XX New isolated human genes and the secreted polypeptides they encode
XX useful for diagnosis and treatment of e.g. cancers, neurological
XX disorders, immune diseases, inflammation or blood disorders
XX
XX Claim 11: Page 313-315; 385pp; English.
XX
XX This sequence represents a secreted human protein encoded by the gene
XX clone detailed in the descriptor line. The gene can be used to generate
XX fusion proteins by linking to the gene to a human immunoglobulin Fc
XX portion (e.g. AAX00602) for increasing the stability of the fused
XX protein as compared to the human protein only.
XX The invention relates to 87 novel genes and their fragments (nucleic
XX acid sequences: AAX00611-X00724; amino acid sequences AAW67807-W68004)
XX which are useful for preventing, treating or ameliorating medical
XX conditions e.g. by protein or gene therapy. Also, pathological
XX conditions can be diagnosed by determining the amount of the new
XX polypeptides in a sample or by determining the presence of mutations in
XX the new polynucleotides. Specific uses are described for each of the 87
XX polynucleotides, based on which tissues they are most highly expressed in
XX (see AAX00611 for described uses).
XX
XX SQ Sequence 625 AA;
Query Match 6.2%; Score 38; DB 20; Length 625;
Best Local Similarity 100.0%; Pred. No. 3e-29;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 YQAHVSRNRKGGVYVTRGGFRGCTVWLTGLSGAGKTT 57
| | | | | | | | | | | | | | | | | | | | | |
Db 30 YQAHVSRNRKGGVYVTRGGFRGCTVWLTGLSGAGKTT 67
| | | | | | | | | | | | | | | | | | | | | |
RESULT 13
AAW75032
ID AAW75032 standard; Protein: 36 AA.
XX
XX AAW75032;
XX
XX 25-JAN-1999 (first entry)
XX
XX Fragment of human secreted protein encoded by gene 14.
XX
XX Human; secreted protein; testis; tumour; foetal brain tissue;
XX fusion protein; cancer; central nervous system; seizure;
XX diagnosis; neurodegenerative disease.
XX
XX Homo sapiens.
XX
XX WO9839448-A2.
XX
XX 11-SEP-1998.
XX
XX 06-MAR-1998; 98WO-US04493.
XX
XX 02-OCT-1997; 97US-0061060.
PR 07-MAR-1997; 97US-0038621.
PR 07-MAR-1997; 97US-0040161.
PR 07-MAR-1997; 97US-0040162.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043589.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 23-MAY-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 13-JUN-1997; 97US-0049610.
PR 08-JUL-1997; 97US-0051926.
PR 16-JUL-1997; 97US-0052874.
PR 18-AUG-1997; 97US-0055724.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
```

PR 22-AUG-1997; 97US-0056875.
 PR 22-AUG-1997; 97US-0056876.
 PR 22-AUG-1997; 97US-0056877.
 PR 22-AUG-1997; 97US-0056878.
 PR 22-AUG-1997; 97US-0056879.
 PR 22-AUG-1997; 97US-0056880.
 PR 22-AUG-1997; 97US-0056881.
 PR 22-AUG-1997; 97US-0056882.
 PR 22-AUG-1997; 97US-0056883.
 PR 22-AUG-1997; 97US-0056884.
 PR 22-AUG-1997; 97US-0056885.
 PR 22-AUG-1997; 97US-0056886.
 PR 22-AUG-1997; 97US-0056887.
 PR 22-AUG-1997; 97US-0056888.
 PR 22-AUG-1997; 97US-0056889.
 PR 22-AUG-1997; 97US-0056892.
 PR 22-AUG-1997; 97US-0056893.
 PR 22-AUG-1997; 97US-0056894.
 PR 22-AUG-1997; 97US-0056903.
 PR 22-AUG-1997; 97US-0056908.
 PR 22-AUG-1997; 97US-0056909.
 PR 22-AUG-1997; 97US-0056910.
 PR 22-AUG-1997; 97US-0056911.
 PR 05-SEP-1997; 97US-0057650.
 PR 05-SEP-1997; 97US-0057669.
 PR 05-SEP-1997; 97US-0057761.
 PR 12-SEP-1997; 97US-0058785.
 XX

(HUMA-) HUMAN GENOME SCI INC.

PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
 PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
 PI Kyaw H, Lafleur DW, Li Y, Moore FA, Ni J, Olsen HS, Rosen CA;
 PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
 XX
 DR WPI; 1998-506364/43.

XX New isolated human genes and the secreted polypeptide(s) they encode
 PT useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 XX

PS Disclosure; Page 14; 721pp; English.

XX This sequence represents a fragment of a secreted human protein encoded
 CC by the nucleic acid molecule designated Gene 14 (AAV59524).
 CC The gene can be used to generate fusion proteins by linking to the gene
 CC to a human immunoglobulin Fc portion (e.g. AAV59502) for increasing the
 CC stability of the fused protein as compared to the human protein only.
 CC The invention relates to 186 novel genes and their fragments (nucleic
 CC acid sequences: AAV59511-V59812; amino acid sequences AAV74731-W75026)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 186
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAV59511 for described uses).

XX Sequence 36 AA;

Query Match 5.9%; Score 36; DB 19; Length 36;

Best Local Similarity 100.0%; Pred. No. 2.5e-28;

Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRQGVVGRGFRGCTVWLTLGSGAGK 55

Db 1 YQAHVSRNKRQGVVGRGFRGCTVWLTLGSGAGK 36

RESULT 14

AAG44041

ID AAG44041 standard; Protein: 373 AA.

XX

AC AAG44041;

XX

DT 18-OCT-2000 (first entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 55117.
 DE
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS
 XX EP1033405-A2.
 DN
 XX 06-SEP-2000.
 PD
 XX 25-FEB-2000; 2000EP-0301439.
 PF
 XX 25-FEB-1999; 99US-0121825.
 XX 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.

PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.6%; Score 22; DB 21; Length 373;
Best Local Similarity 100.0%; Pred. No. 2.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAFOLRNPVNHGALLM 422
Db 149 NADAVFAFOLRNPVNHGALLM 170

RESULT 15

AAG44040

ID AAG44040 standard; Protein; 460 AA.

XX AAG44040;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55116.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000. XX
XX PF 25-FEB-2000; 2000EP-0301439.
XX XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.

PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.6%; Score: 22; DB 21; Length 460;

Best Local Similarity 100.0%; Pred. No. 3.4e-13; Mismatches 0; Indels 0; Gaps 0;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAQLRNPVHNGHALLM 422

|||||

Db 236 NADAVFAQLRNPVHNGHALLM 257

Search completed: May 30, 2003, 12:30:13
Job time : 75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:29:00 ; Search time 27 Seconds
(without alignments)
670.189 Million cell updates/sec

Title: US-09-898-165B-7
Perfect score: 615
Sequence: 1 MSGIKKOTENQOKSTNVY.....MAPKAWKVLTDYRSEMDKN 615

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 10

Total number of hits satisfying chosen parameters: 14

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	8.9	174	4	US-09-149-476-635
2	38	6.2	624	2	US-08-879-561-3
3	38	6.2	624	2	US-08-879-561-9
4	36	5.9	36	4	US-09-149-476-634
5	22	3.6	110	4	US-09-346-408-10
6	22	3.6	133	4	US-09-346-408-11
7	22	3.6	463	4	US-09-346-408-11
8	22	3.6	465	4	US-09-346-408-4
9	22	3.6	465	4	US-09-346-408-6
10	22	3.6	610	2	US-08-879-561-10
11	21	3.4	461	4	US-09-346-408-8
12	21	3.4	476	4	US-09-346-408-12
13	12	2.0	27	4	US-09-149-476-637
14	11	1.8	204	4	US-09-134-001C-4517

ALIGNMENTS

RESULT 1
US-09-149-476-635
; Sequence 635, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1

; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671

EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 8.9%; Score 55; DB 4: Length 174;

Best Local Similarity 100.0%; Pred. No. 6.le-46; Mismatches 0; Indels 0; Gaps 0;

Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGIKKQKTENQKSTNVVYQAHHSNRKRGQVVGTRGGFRGCTVWLTLGSGAGK 55

Db 55 MSGIKKQKTENQKSTNVVYQAHHSNRKRGQVVGTRGGFRGCTVWLTLGSGAGK 109

RESULT 2

US-08-879-561-3

; Sequence 3, Application US/08879561

; Patent No. 5817482

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Hawkins, Phillip R.

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/879,561
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0325 US
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGN0202
CLONE: 373887
US-08-879-561-3

Query Match 6.2%; Score 38; DB 2; Length 624;
Best Local Similarity 100.0%; Pred. No. 8e-29;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 YQAHVSRNKRQVVGTRGFRGCTVWLTGLSGAGKTT 57
|||||
Db 30 YQAHVSRNKRQVVGTRGFRGCTVWLTGLSGAGKTT 67

RESULT 3
US-08-879-561-9
Sequence 9, Application US/08879561
Patent No. 5817482
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Hawkins, Phillip R.
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/879,561
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:

APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0325 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1109676
US-08-879-561-9

Query Match 6.2%; Score 38; DB 2; Length 624;
Best Local Similarity 100.0%; Pred. No. 8e-29;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 YQAHVSRNKRQVVGTRGFRGCTVWLTGLSGAGKTT 57
|||||
Db 30 YQAHVSRNKRQVVGTRGFRGCTVWLTGLSGAGKTT 67

RESULT 4
US-09-149-476-634
Sequence 634, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503

[illegible]

```

; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02
;

```

Query Match 5.9%; Score 36; DB 4; Length 36;
Best Local Similarity 100.0%; Pred. No. 6.3e-28;
Matches 36; Conservative 0; Mismatches 0; Indels

QY	20	YQAHHVSRNKGQVVGTRGGFRCCTVWLTLGLSGAGK	55
Db	1	YQAHHVSRNKGQVVGTRGGFRCCTVWLTLGLSGAGK	36

RESULT 5

US-09-346-408-10
; Sequence 10, Application US/09346408B
; Patent No. 6338966

; ACCENT NO: 00502000
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Anderson, Shawn
 ; APPLICANT: Falco, Carl
 ; APPLICANT: Rafalski, Antoni

```

> > INVENTOR: Genes Encoding Sulfate Assimilation Proteins
> TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
> FILE REFERENCE: BB-1167-A
> CURRENT APPLICATION NUMBER: US/09/346,408B
> CURRENT FILING DATE: 1999-07-01
> EARLIER APPLICATION NUMBER: 60/092,833
> EARLIER FILING DATE: July 14, 1998
> NUMBER OF SEQ ID NOS: 12
> SOFTWARE: Microsoft Office 97
> SEQ ID NO 10

```

; ORGANISM: TYP
IIS-09-346-408-10

```
Query Match      3.6%; Score 22; DB 4; Length 110;
Best Local Similarity 100.0%; Pred. No. 8.3e-14;
Matches 22; Conservative 0; Mismatches 0; Indels
```

QY 401 NADAVFAQLRNPVHGHALLM 422
 |||||
 Db 45 NADAVFAQLRNPVHGHALLM 66

RESULT 6

US-09-346-408-2
; Sequence 2, Application US/09346408B
; Patent No. 6338966

APPLICANT: Allen, Steve
APPLICANT: Anderson, Sh
APPLICANT: Falco, Carl
APPLICANT: Rafalski, An

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

; LENGTH: 133
; TYPE: PRT
; ORGANISM: Z
US-09-346-408-2

```

Query Match 3.6%; Score 22; DB 4; Length 133;
Best Local Similarity 100.0%; Pred. No. 9.8e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	401	NADAVFAQLRNPVHNGHALLM	422
Db	21	NADAVFAQLRNPVHNGHALLM	42

RESULT 7

US-09-346-408-11
; Sequence 11, Application US/09346408B
; Patent No. 6338966

GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Anderson, Shawn
APPLICANT: Falco, Carl
APPLICANT: Rafalski, Anton

```

1  TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
2
3  FILE REFERENCE: BB-1167-A
4
5  CURRENT APPLICATION NUMBER: US/09/346,408B
6
7  CURRENT FILING DATE: 1999-07-01
8
9  EARLIER APPLICATION NUMBER: 60/092,833
10
11 EARLIER FILING DATE: July 14, 1998
12
13 NUMBER OF SEQ ID NOS: 12
14
15 SOFTWARE: Microsoft Office 97
16
17 SEQ ID NO 11
18
19 LENGTH: 463

```

US-09-346-408-11

Query Match 3.6%; Score 22; DB 4; Length 463;
Best Local Similarity 100.0%; Pred.No. 2.8e-13;
Matches 22; Conservative 0; Mismatches 0; Indels

Qy	401	NADAVFAQLRNPVHNGHALLM	422
Db	239	NADAVFAQLRNPVHNGHALLM	260

RESULT 8

US-09-346-408-4
; Sequence 4, Application US/09346408B
; Patent No. 6338966

;
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falco, Carl
; APPLICANT: Rafalski, Antoni

```

? COUNTRY: UNKNOWN
? TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
? FILE REFERENCE: BB-1167-A
? CURRENT APPLICATION NUMBER: US/09/346,408B
? CURRENT FILING DATE: 1999-07-01
? EARLIER APPLICATION NUMBER: 60/092,833
? EARLIER FILING DATE: July 14, 1998
? NUMBER OF SEQ ID NOS: 12
? SOFTWARE: Microsoft Office 97
? SEQ ID NO 4
? LENGTH: 465

```

; ORGANISM: M
US-09-346-408-4

Query Match 3.6%; Score 22; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.9e-13;
Matches 22: Conservative 0; Mismatches 0; Indels

; SEQ ID NO 2

QY 401 NADAVFAQLRNPVHNGHALLM 422
|||||
DB 241 NADAVFAQLRNPVHNGHALLM 262

RESULT 9

US-09-346-408-6
; Sequence 6, Application US/09346408B
; Patent No. 6338966

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falco, Carl
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-A
; CURRENT APPLICATION NUMBER: US/09/346,408B
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 60/092,833
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6

; LENGTH: 465

; TYPE: PRT

; ORGANISM: Glycine max

US-09-346-408-6

Query Match 3.6%; Score 22; DB 4; Length 465;
Best Local Similarity 100.0%; Pred. No. 2.9e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAQLRNPVHNGHALLM 422
|||||
DB 241 NADAVFAQLRNPVHNGHALLM 262

RESULT 10

US-08-879-561-10
; Sequence 10, Application US/08879561
; Patent No. 5817482

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA: US/08/879,561

; APPLICATION NUMBER: 424

; FILING DATE: Herewith

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0325 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:

; INFORMATION FOR SEQ ID NO: 10:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 610 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 705385

US-08-879-561-10

Query Match 3.6%; Score 22; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 3.6e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAQLRNPVHNGHALLM 422
|||||
DB 396 NADAVFAQLRNPVHNGHALLM 417

RESULT 11

US-09-346-408-8

; Sequence 8, Application US/09346408B

; Patent No. 6338966

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Anderson, Shawn

; APPLICANT: Falco, Carl

; APPLICANT: Rafalski, Antoni

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167-A

; CURRENT APPLICATION NUMBER: US/09/346,408B

; CURRENT FILING DATE: 1999-07-01

; EARLIER APPLICATION NUMBER: 60/092,833

; EARLIER FILING DATE: July 14, 1998

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 8

; LENGTH: 461

; TYPE: PRT

; ORGANISM: Triticum aestivum

US-09-346-408-8

Query Match 3.4%; Score 21; DB 4; Length 461;
Best Local Similarity 100.0%; Pred. No. 2.7e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 402 ADAVFAQLRNPVHNGHALLM 422
|||||
DB 239 ADAVFAQLRNPVHNGHALLM 259

RESULT 12

US-09-346-408-12

; Sequence 12, Application US/09346408B

; Patent No. 6338966

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Anderson, Shawn

; APPLICANT: Falco, Carl

; APPLICANT: Rafalski, Antoni

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167-A

; CURRENT APPLICATION NUMBER: US/09/346,408B

; CURRENT FILING DATE: 1999-07-01

; EARLIER APPLICATION NUMBER: 60/092,833

; EARLIER FILING DATE: July 14, 1998

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 12

```
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-346-408-12

Query Match          3.4%; Score 21; DB 4; Length 476;
Best Local Similarity 100.0%; Pred. No. 2.8e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 402 ADAVFAFQLRNPVNHGHALLM 422
DB 254 ADAVFAFQLRNPVNHGHALLM 274

RESULT 13
; Sequence 637, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
```

EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 2.0%; Score 12; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 0.00015;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 LDGDNVRHGLNR 87
|||||
Db 2 LDGDNVRHGLNR 13
RESULT 14
US-09-134-001C-4517
; Sequence 4517, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4517
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4517

Query Match 1.8%; Score 11; DB 4; Length 204;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LDGDNVRHGLN 86
|||||
Db 66 LDGDNVRHGLN 76

Search completed: May 30, 2003, 12:33:43
Job time : 28 secs

GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 39981
LENGTH: 65
TYPE: PRT
ORGANISM: Homo sapiens
OTHER INFORMATION: MAP TO AC006191.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
OTHER INFORMATION: SWISSPROT HIT: O95340, EVALUATE 4.00e-33
OTHER INFORMATION: EST_HUMAN HIT: BE292722.1, EVALUATE 4.00e-32

US-09-864-761-39981

Query Match 10.6%; Score 65; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.6e-55;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EYLVSHATPCVSLDGNVHGLNRLNLFSPGDREENIRRIAEVAKLFADAGLVCTTST 122
|||||
DB 1 EYLVSHATPCVSLDGNVHGLNRLNLFSPGDREENIRRIAEVAKLFADAGLVCTTST 60
|||||

QY 123 SPFAK 127
|||||
DB 61 SPFAK 65

RESULT 3
US-09-809-391-635
Sequence 635, Application US/09809391
Publication No. US20030049618A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P2
CURRENT APPLICATION NUMBER: US/09/809,391
CURRENT FILING DATE: 2001-03-16
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 761
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 635
LENGTH: 174
TYPE: PRT
ORGANISM: Homo sapiens
US-09-809-391-635

Query Match 8.9%; Score 55; DB 9; Length 174;
Best Local Similarity 100.0%; Pred. No. 2e-45;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGIKKQKTENQOKSTNVVYQAHVSRNKRQGVVGTGRGCTVWLTLGSLGAGK 55
|||||
DB 55 MSGIKKQKTENQOKSTNVVYQAHVSRNKRQGVVGTGRGCTVWLTLGSLGAGK 109
|||||

RESULT 4
US-09-984-245-196
Sequence 196, Application US/09984245
Patent No. US20020165374A1
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 87 Human Secreted Proteins
FILE REFERENCE: P2004P1
CURRENT APPLICATION NUMBER: US/09/984,245
CURRENT FILING DATE: 2001-10-29
Prior APPLICATION NUMBER: 09/154,707
Prior FILING DATE: 1998-09-17
Prior APPLICATION NUMBER: PCT/US98/05311
Prior FILING DATE: 1998-03-19
Prior APPLICATION NUMBER: US 60/041,277
Prior FILING DATE: 1997-03-21
Prior APPLICATION NUMBER: US 60/042,344
Prior FILING DATE: 1997-03-21
Prior APPLICATION NUMBER: US 60/041,276
Prior FILING DATE: 1997-03-21
Prior APPLICATION NUMBER: US 60/041,281
Prior FILING DATE: 1997-03-21
Prior APPLICATION NUMBER: US 60/048,094
Prior FILING DATE: 1997-05-30
Prior APPLICATION NUMBER: US 60/048,350
Prior FILING DATE: 1997-05-30
Prior APPLICATION NUMBER: US 60/048,188
Prior FILING DATE: 1997-05-30
Prior APPLICATION NUMBER: US 60/048,135
Prior FILING DATE: 1997-05-30
Prior APPLICATION NUMBER: US 60/050,937
Prior FILING DATE: 1997-05-30
Prior APPLICATION NUMBER: US 60/048,187
Prior FILING DATE: 1997-05-30
Prior APPLICATION NUMBER: US 60/048,099
Prior FILING DATE: 1997-05-30
Prior APPLICATION NUMBER: US 60/048,352
Prior FILING DATE: 1997-05-30
Prior APPLICATION NUMBER: US 60/048,186
Prior FILING DATE: 1997-05-30

RESULT 5
US-09-366-262-196
; Sequence 196, Application US/09966262
; Publication No. US20030050461A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Pro
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/966,262
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30

RESULT 6

US-09-983-966-196
; Sequence 196, Application US/09983966
; Publication No. US20030060619A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/983,966
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30

;; PRIOR APPLICATION NUMBER: US 60/048,135
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/050,937
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,187
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,099
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,352
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,186
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,069
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,095
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,131
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,096
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,355
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/054,804
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: US 60/056,370
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: US 60/060,862
;; PRIOR FILING DATE: 1997-10-02
;; NUMBER OF SEQ ID NOS: 343
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 196
;; LENGTH: 624
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-983-966-196

Query Match 6.2%; Score 38; DB 9; Length 624;
Best Local Similarity 100.0%; Pred. No. 2.1e-28;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRKGQVGTGRGCTVWLTGLSGAGKTT 57
|||||
DB 30 YQAHVSRNKRKGQVGTGRGCTVWLTGLSGAGKTT 67

RESULT 7
US-10-143-090-196
;; Sequence 196, Application US/10143090
;; Publication No. US20030069406A1
;; GENERAL INFORMATION:
;; APPLICANT: Young et al.
;; TITLE OF INVENTION: 87 Human Secreted Proteins
;; FILE REFERENCE: P2004P1
;; CURRENT APPLICATION NUMBER: US/10/143,090
;; CURRENT FILING DATE: 2002-05-13
;; PRIOR APPLICATION NUMBER: 09/154,707
;; PRIOR FILING DATE: 1998-09-17
;; NUMBER OF SEQ ID NOS: 343
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 196
;; LENGTH: 624
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-143-090-196

Query Match 6.2%; Score 38; DB 9; Length 624;
Best Local Similarity 100.0%; Pred. No. 2.1e-28;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 YQAHVSRNKRKGQVGTGRGCTVWLTGLSGAGKTT 57
|||||
DB 30 YQAHVSRNKRKGQVGTGRGCTVWLTGLSGAGKTT 67

RESULT 8
US-09-771-161A-244
;; Sequence 244, Application US/09771161A
;; Patent No. US20020110811A1
;; GENERAL INFORMATION:
;; APPLICANT: LEVINE, et al.
;; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
;; FILE REFERENCE: 802620-2005.1
;; CURRENT APPLICATION NUMBER: US/09/771.161A
;; CURRENT FILING DATE: 2001-01-26
;; PRIOR APPLICATION NUMBER: 09/724,676
;; PRIOR FILING DATE: 2000-11-28
;; PRIOR APPLICATION NUMBER: 136776
;; PRIOR FILING DATE: 2000-06-15
;; PRIOR APPLICATION NUMBER: 135619
;; PRIOR FILING DATE: 2000-04-12
;; NUMBER OF SEQ ID NOS: 273
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 244
;; LENGTH: 624
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-771-161A-244

Query Match 6.2%; Score 38; DB 10; Length 624;
Best Local Similarity 100.0%; Pred. No. 2.1e-28;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRKGQVGTGRGCTVWLTGLSGAGKTT 57
|||||
DB 30 YQAHVSRNKRKGQVGTGRGCTVWLTGLSGAGKTT 67

RESULT 9
US-09-809-391-634
;; Sequence 634, Application US/09809391
;; Publication No. US20030049618A1
;; GENERAL INFORMATION:
;; APPLICANT: Ruben et al.
;; TITLE OF INVENTION: 186 Human Secreted proteins
;; FILE REFERENCE: P2002P2
;; CURRENT APPLICATION NUMBER: US/09/809,391
;; CURRENT FILING DATE: 2001-03-16
;; Prior application data removed - consult PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 761
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 634
;; LENGTH: 36
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-809-391-634

Query Match 5.9%; Score 36; DB 9; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.6e-27;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRKGQVGTGRGCTVWLTGLSGAGK 55
|||||
DB 1 YQAHVSRNKRKGQVGTGRGCTVWLTGLSGAGK 36

RESULT 10
US-09-984-245-332
;; Sequence 332, Application US/09984245
;; Patent No. US20020165374A1
;; GENERAL INFORMATION:

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 340
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-245-340

Query Match 3.3%; Score 20; DB 9; Length 41;
Best Local Similarity 100.0%; Pred. No. 6.7e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 AGANFYIVGRDPAGMPHPET 522
Db 18 AGANFYIVGRDPAGMPHPET 37

RESULT 15

US-09-966-262-340
; Sequence 340, Application US/09966262
; Publication No. US20030050461A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004p1
; CURRENT APPLICATION NUMBER: US/09/966,262
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350

; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 340
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-262-340

Query Match 3.3%; Score 20; DB 9; Length 41;
Best Local Similarity 100.0%; Pred. No. 6.7e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 AGANFYIVGRDPAGMPHPET 522
Db 18 AGANFYIVGRDPAGMPHPET 37

Search completed: May 30, 2003, 12:41:34
Job time : 50 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:28:05 ; Search time 43 Seconds
(without alignments)
1374.946 Million cell updates/sec

Title: US-09-898-165B-7
Perfect score: 615
Sequence: 1 MSGIKKKTENQOKSTNVY.....MAPKAWLVLDYRSEMDKN 615

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 282224 seqs, 96134422 residues
Word size : 10

Total number of hits satisfying chosen parameters: 32

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR-73.*

- 1: pir1.*
- 2: pir2.*
- 3: pir3.*
- 4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	6.2	624	1 JW0087	adenyl-yl-sulfate k
2	22	3.6	424	2 S44079	sulfate adenyl-yltr
3	22	3.6	463	2 S44267	sulfate adenyl-yltr
4	22	3.6	489	1 T01204	sulfate adenyl-yltr
5	22	3.6	610	1 J04383	adenyl-yl-sulfate k
6	21	3.4	459	2 T08594	probable sulfate a
7	21	3.4	465	2 E71409	sulfate adenyl-yltr
8	21	3.4	469	2 T52659	sulfate adenyl-yltr
9	21	3.4	476	2 S44943	sulfate adenyl-yltr
10	21	3.4	483	2 T14475	sulfate adenyl-yltr
11	18	2.9	652	1 T24918	3'-phosphoadenosin
12	13	2.1	177	1 S74917	adenyl-yl-sulfate k
13	13	2.1	202	2 T50101	adenyl-yl-sulfate k
14	12	2.0	276	1 S47640	adenyl-yl-sulfate k
15	12	2.0	293	2 T06100	adenyl-yl-sulfate k
16	12	2.0	633	2 H95932	probable adenyl-yl-
17	12	2.0	641	1 Z22RNQ	adenyl-yl-sulfate k
18	12	2.0	641	2 E95320	adenyl-yl-sulfate k
19	12	2.0	660	2 G82672	ATP sulfurylase, l
20	11	1.8	110	2 C55228	hypothetical prote
21	11	1.8	196	2 H83472	adenosine 5'-phosp
22	11	1.8	214	1 S18729	adenyl-yl-sulfate k
23	11	1.8	312	2 T08076	adenyl-yl-sulfate k
24	11	1.8	635	2 A87433	hypothetical prote
25	11	1.8	644	2 AD3471	adenyl-yl-sulfate k
26	10	1.6	200	2 E96912	adenyl-yl-sulfate k
27	10	1.6	201	1 B65056	adenyl-yl-sulfate k
28	10	1.6	201	2 D91079	adenosine 5'-phosp
29	10	1.6	201	2 E85924	adenosine 5'-phosp

30 10 1.6 201 2 AH0856 adenosine 5-phosph
31 10 1.6 213 2 AF0408 adenyl-yl-sulfate k
32 10 1.6 215 2 F82062 adenyl-yl-sulfate k

ALIGNMENTS

RESULT 1
JW0087

adenyl-yl-sulfate kinase (EC 2.7.1.25) - human
N:Alternate names: adenosine 5'-phosphosulfate kinase: PAPS
N:Contains: adenylsulfate kinase (EC 2.7.1.25); sulfate adenyltransferase (EC 2.7.1.25)
C:Species: Homo sapiens (man)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002
C:Accession: JW0087
R:Yanagisawa, K.; Sakakibara, Y.; Suiko, M.; Takami, Y.; Nakayama, T.; Nakajima, H.; Biosci. Biotechnol. Biochem. 62, 1037-1040, 1998
A:Title: cDNA cloning, expression, and characterization of the human bifunctional A.P
A:Reference number: JW0087; MUID:98312048; PMID:9648242
A:Accession: JW0087
A:Molecule type: mRNA
A:Residues: 1-624 <YAN>
A:Cross-references: GB:AF033026; NID:g33378100; PIDN:AAC28429.1; PID:g33378101
A:Experimental source: Brain
C:Function: <ASKF>
A:Description: as adenylsulfate kinase catalyzes the phosphorylation of adenyl-ylsulfate
C:Function: <SATF>
A:Description: as sulfate adenyltransferase catalyzes the reaction of sulfate and A
C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenylsulfate
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loop
F:52-215/Domain: adenylsulfate kinase homology <ASK>
F:59-66/Region: nucleotide-binding motif A (P-loop)
F:226-620/Domain: sulfate adenyltransferase homology <SAT>
F:133/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 6.2%; Score 38; DB 1; Length 624;
Best Local Similarity 100.0%; Pred. No. 2e-30; Indels 0; Gaps 0;
Matches 38; Conservative 0; Mismatches 0;

QY 20 YQAHVSRNKRGOVVGTRGGRGCTVWLTGLSGAGKTT 57
|||||
DB 30 YQAHVSRNKRGOVVGTRGGRGCTVWLTGLSGAGKTT 67

RESULT 2
S44079

sulfate adenyltransferase (EC 2.7.7.4) met3-1 - potato
C:Species: Solanum tuberosum (potato)
C:Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C:Accession: S67497; S44079
R:Klonus, D.; Hoefgen, R.; Willmitzer, L.; Riesmeier, J.W.
Plant J. 6, 105-112, 1994
A:Title: Isolation and characterization of two cDNA clones encoding ATP-sulfurylase
A:Reference number: S67497; MUID:95004649; PMID:7920699
A:Accession: S67497
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-424 <KLO>
A:Cross-references: EMBL:X75041; NID:g471342; PIDN:CAA52953.1; PID:g531495
C:Genetics: met3-1
A:Gene: met3-1
C:Superfamily: sulfate adenyltransferase met3-1; sulfate adenyltransferase homo-
C:Keywords: nucleotidyltransferase
F:11-409/Domain: sulfate adenyltransferase homology <SAT>

Query Match 3.6%; Score 22; DB 2; Length 424;
Best Local Similarity 100.0%; Pred. No. 4.6e-14; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 401 NADAVFAFQLRNPVINGHALLM 422
|||||
DB 200 NADAVFAFQLRNPVINGHALLM 221

Gene 165, 243-248, 1995
A:Title: A multifunctional Urechis caupo protein, PAPS synthetase, has both ATP sulfa
A:Reference number: JC4383; MUID:96096529; PMID:8522184
A:Accession: JC4383
A:Molecule type: mRNA
A:Residues: 1-610 <ROS>
A:Cross-references: GB:L39001; NID:g705384; PIDN:AAB00139.1; PID:g705385
C:Genetics:
A:Gene: papss
C:Function: <ASK>
A:Description: as adenylylsulfate kinase catalyzes the phosphorylation of adenylylsul
C:Function: <SAT>
A:Description: as sulfate adenylyltransferase catalyzes the reaction of sulfate and A
C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenylylsulfa
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loo
F:37-200/Domain: adenylylsulfate kinase homology <ASK>
F:44-51/Region: nucleotide-binding motif A (P-loop)
F:161-171/Domain: 3'-phosphoadenosine-5'-phosphosulfate binding #status predicted <PA
F:211-605/Domain: sulfate adenylyltransferase homology <SAT>
F:118/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 3.6%; Score 22; DB 1; Length 610;
Best Local Similarity 100.0%; Pred. No. 6.5e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVFAFQLRNPVHNGHALLM 422
|||||
Db 396 NADAVFAFQLRNPVHNGHALLM 417
|||||

RESULT 6
T08594
probable sulfate adenylyltransferase (EC 2.7.7.4) - rape
C:Species: Brassica napus (rape)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08594
R:Buchanan-Wollaston, V.; Ainsworth, C.
Plant Mol. Biol. 33, 821-834, 1997
A:Title: Leaf senescence in Brassica napus: cloning of senescence related genes by su
A:Reference number: Z16446; MUID:97260386; PMID:9106506
A:Accession: T08594
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-459 <BUC>
A:Cross-references: EMBL:U68218; NID:g1527218; PIDN:AAB53100.1; PID:g1527219
A:Experimental source: cv. Falcon; senescing leaves
C:Genetics:
A:Gene: LSC680
C:Function:
A:Description: catalyzes the activation of sulfate to adenylylsulfate
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homolo
C:Keywords: nucleotidyltransferase
F:45-444/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 3.4%; Score 21; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 ADAVFAFQLRNPVHNGHALLM 422
|||||
Db 236 ADAVFAFQLRNPVHNGHALLM 256
|||||

RESULT 7
E71409
sulfate adenylyltransferase (EC 2.7.7.4) precursor (clone APS3) - Arabidopsis thalian
N:Alternate names: ATP sulfurylase
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
C:Accession: E71409; S68201
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; G

Gene 165, 243-248, 1995
A:Title: A multifunctional Urechis caupo protein, PAPS synthetase, has both ATP sulfa
A:Reference number: JC4383; MUID:96096529; PMID:8522184
A:Accession: JC4383
A:Molecule type: mRNA
A:Residues: 1-610 <ROS>
A:Cross-references: GB:L39001; NID:g705384; PIDN:AAB00139.1; PID:g705385
C:Genetics:
A:Gene: papss
C:Function: <ASK>
A:Description: as adenylylsulfate kinase catalyzes the phosphorylation of adenylylsul
C:Function: <SAT>
A:Description: as sulfate adenylyltransferase catalyzes the reaction of sulfate and A
C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenylylsulfa
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loo
F:37-200/Domain: adenylylsulfate kinase homology <ASK>
F:44-51/Region: nucleotide-binding motif A (P-loop)
F:161-171/Domain: 3'-phosphoadenosine-5'-phosphosulfate binding #status predicted <PA
F:211-605/Domain: sulfate adenylyltransferase homology <SAT>
F:118/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 3.6%; Score 22; DB 1; Length 610;
Best Local Similarity 100.0%; Pred. No. 6.5e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVFAFQLRNPVHNGHALLM 422
|||||
Db 396 NADAVFAFQLRNPVHNGHALLM 417
|||||

RESULT 6
T08594
probable sulfate adenylyltransferase (EC 2.7.7.4) - rape
C:Species: Brassica napus (rape)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08594
R:Buchanan-Wollaston, V.; Ainsworth, C.
Plant Mol. Biol. 33, 821-834, 1997
A:Title: Leaf senescence in Brassica napus: cloning of senescence related genes by su
A:Reference number: Z16446; MUID:97260386; PMID:9106506
A:Accession: T08594
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-459 <BUC>
A:Cross-references: EMBL:U68218; NID:g1527218; PIDN:AAB53100.1; PID:g1527219
A:Experimental source: cv. Falcon; senescing leaves
C:Genetics:
A:Gene: LSC680
C:Function:
A:Description: catalyzes the activation of sulfate to adenylylsulfate
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homolo
C:Keywords: nucleotidyltransferase
F:45-444/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 3.4%; Score 21; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 ADAVFAFQLRNPVHNGHALLM 422
|||||
Db 236 ADAVFAFQLRNPVHNGHALLM 256
|||||

RESULT 7
E71409
sulfate adenylyltransferase (EC 2.7.7.4) precursor (clone APS3) - Arabidopsis thalian
N:Alternate names: ATP sulfurylase
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
C:Accession: E71409; S68201
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; G

avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalwatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number: A71400; MUID:96121113; PMID:9461215
A:Accession: E71409
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-465 <BEV>
A:Cross-references: GB:297336; NID:92244788; PIDN:CAB10247.1; PID:92244824
R:Murillo, M.; Leustek, T.
Arch. Biochem. Biophys. 323, 195-204, 1995
A:Title: Adenosine-5'-triphosphate-sulfurylase from Arabidopsis thaliana and Escherichia
ine-5'-triphosphate-sulfurylase cDNAs from Arabidopsis thaliana and analysis of a recom
A:Reference number: S68024; MUID:96019964; PMID:7487067
A:Accession: S68201
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-6, 'D', 8-35, 'A', 37-465 <MUR>
A:Cross-references: EMBL:U06275; NID:9459143; PIDN:AAA923350.1; PID:9459144
C:Genetics:
A:Map position: 4COP9-4G3845
A:Genome: nuclear
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology
C:Keywords: chloroplast; nucleotidyltransferase
F:1-50/Domain: transit peptide (chloroplast) #status predicted <TNP>
F:51-465/Product: sulfate adenylyltransferase #status predicted <MAT>
F:52-450/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 3.4%; Score 21; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 ADAVFATQLRNPVHNGHALLM 422
Db 242 ADAVFATQLRNPVHNGHALLM 262
|||||

RESULT 8
T52659
sulfate adenylyltransferase (EC 2.7.7.4) ap54 precursor [validated] - Arabidopsis thalia
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 03-Nov-2000
C:Accession: T52659
R:Hatzfeld, Y.; Lee, S.; Lee, M.; Leustek, T.; Saito, K.
Gene 248, 51-58, 2000
A:Title: Functional characterization of a gene encoding a fourth ATP sulfurylase isoform
A:Reference number: Z26165
A:Accession: T52659
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-469 <HAT>
A:Cross-references: EMBL:AJ012586; PIDN:CAB42640.1
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ap54
A:Genome: nuclear
C:Function:
A:Description: EC 2.7.7.4 [validated, MUID:20267847]
A:Pathway: sulfate assimilation
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology
C:Keywords: chloroplast; nucleotidyltransferase

Query Match 3.4%; Score 21; DB 2; Length 469;
Best Local Similarity 100.0%; Pred. No. 5.5e-13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 ADAVFATQLRNPVHNGHALLM 422
Db 244 ADAVFATQLRNPVHNGHALLM 264
|||||

RESULT 9

S44943
sulfate adenylyltransferase (EC 2.7.7.4) met3-1 precursor - Arabidopsis thaliana
N:Alternate names: ATP sulfurylase; sulfate adenylyltransferase APS2
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 06-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C:Accession: S44943; S68024; S68202
R:Klonus, D.
submitted to the EMBL Data Library, May 1994
A:Description: ATP-sulfurylase cDNA clone from A.thaliana.
A:Reference number: S44943
A:Accession: S44943
A:Molecule type: mRNA
A:Residues: 1-476 <KLO>
A:Cross-references: EMBL:X79210; NID:9487403; PIDN:CAA55799.1; PID:9487404
R:Murillo, M.; Leustek, T.
Arch. Biochem. Biophys. 323, 195-204, 1995
A:Title: Adenosine-5'-triphosphate-sulfurylase from Arabidopsis thaliana and Escherich
ine-5'-triphosphate-sulfurylase cDNAs from Arabidopsis thaliana and analysis of a rec
A:Reference number: S68024; MUID:96019964; PMID:7487067
A:Accession: S68024
A:Molecule type: mRNA
A:Residues: 1-476 <MUW>
A:Cross-references: EMBL:U06276; NID:91228103; PIDN:AAA92351.1; PID:91228104
A:Experimental source: clone APS2
C:Genetics:
A:Gene: met3-1
A:Genome: nuclear
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homolo
C:Keywords: chloroplast; nucleotidyltransferase
F:1-62/Domain: transit peptide (chloroplast) #status predicted <TNP>
F:63-476/Product: sulfate adenylyltransferase met3-1 #status predicted <MAT>
F:64-462/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 3.4%; Score 21; DB 2; Length 476;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 ADAVFATQLRNPVHNGHALLM 422
Db 254 ADAVFATQLRNPVHNGHALLM 274
|||||

RESULT 10
T14475
sulfate adenylyltransferase (EC 2.7.7.4) ASBo precursor - wild cabbage
N:Alternate names: ATP sulfurylase
C:Species: Brassica oleracea (wild cabbage)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jan-2000
C:Accession: T14475
R:Hatzfeld, Y.; Logan, H.M.; Cathala, N.; Davidian, J.C.
submitted to the EMBL Data Library, September 1996
A:Reference number: Z18111
A:Accession: T14475
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-483 <HAT>
A:Cross-references: EMBL:U69694; NID:91589912; PID:91589913
C:Genetics:
A:Gene: ASBo
A:Introns: 185/2; 407/3; 436/3
C:Function:
A:Description: catalyzes the activation of sulfate to adenylylsulfate
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homolo
C:Keywords: nucleotidyltransferase
F:68-466/Domain: sulfate adenylyltransferase homology <STR>

Query Match 3.4%; Score 21; DB 2; Length 483;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 402 ADAVFATQLRNPVHNGHALLM 422

```
|||||
258 ADAVAFQLRPNVHNGHALLM 278

Db
RESULT 11
T24918
3'-phosphoadenosine-5'-phosphosulfate synthetase - Caenorhabditis elegans
N:Alternate names: protein T14G10.1
C:Species: Caenorhabditis elegans
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 19-Jan-2001
C:Accession: T24918
R:Wild, A.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z19554
A:Accession: T24918
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-652 <WTL>
A:Cross-references: EMBL:Z68880; PIDN:CAA93098.1; GSPDB:GN00022; CESP:T14G10.1
A:Experimental source: clone T14G10.
C:Genetics:
A:Gene: CESP:T14G10.1
A:Map position: 4
A:Introns: 23/3; 82/3; 176/1; 535/2; 623/2
C:Function: <ASK>
A:Description: as adenylylsulfate kinase catalyzes the phosphorylation of adenylylsulfate
C:Function: <SAFE>
A:Description: as sulfate adenylyltransferase catalyzes the reaction of sulfate and ATP
C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenylylsulfate
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loop;
F:56-219/Domain: adenylylsulfate kinase homology <ASK>
F:63-70/Region: nucleotide-binding motif A (P-loop)
F:180-190/Domain: 3'-phosphoadenosine-5'-phosphosulfate binding #status predicted <PAB>
F:230-847/Domain: sulfate adenylyltransferase homology <SAT>
F:137/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 2.9%; Score 18; DB 1; Length 652;
Best Local Similarity 100.0%; Pred. No. 9.3e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 439 LLLHPLGGWTKDDVPLD 456
|||||
Db 476 LLLHPLGGWTKDDVPLD 493

RESULT 12
S74917
adenylyl-sulfate kinase (EC 2.7.1.25) - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein slr0676
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: S74917
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74917
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-177 <KAN>
A:Cross-references: EMBL:D90902; GB:AB001339; NID:g1652027; PIDN:BAAL16957.1; PID:g165203
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: cysC
C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology
C:Keywords: ATP; P-loop; phosphotransferase; purine nucleotide binding
F:5-164/Domain: adenylylsulfate kinase homology <ASK>
F:12-19/Region: nucleotide-binding motif A (P-loop)
. Query Match 2.1%; Score 13; DB 1; Length 177;
```

```
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 WLTGLSGAGKTTI 58
|||||
Db 9 WLTGLSGAGKTTI 21

RESULT 13
T50101
adenylylsulfate kinase [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C:Accession: T50101
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Harris, D.
submitted to the EMBL Data Library, February 2000
A:Reference number: Z25037
A:Accession: T50101
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-202 <WOO>
A:Cross-references: EMBL:AL158056; PIDN:CAB76273.1; GSPDB:GN00066; SPDB:SPAC1782.11
A:Experimental source: strain 972h(-); cosmid cl782
C:Genetics:
A:Gene: SPDB:SPAC1782.11
A:Map position: 1
A:Introns: 157/1
C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology

Query Match 2.1%; Score 13; DB 2; Length 202;
Best Local Similarity 100.0%; Pred. No. 4.5e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 KGLYKRARAGEIK 173
|||||
Db 144 KGLYKRARAGEIK 156

RESULT 14
S47640
adenylyl-sulfate kinase (EC 2.7.1.25) precursor - Arabidopsis thaliana
N:Alternate names: APS kinase; protein F26C24.11
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C:Accession: S47640; T02601; A84521; S38587
R:Arz, H.E.; Gisseimann, G.; Schifmann, S.; Schwenk, J.D.
Biochim. Biophys. Acta 1218, 447-452, 1994
A:Title: A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana.
A:Reference number: S47640; MUID:94325358; PMID:8049272
A:Accession: S47640
A:Molecule type: mRNA
A:Residues: 1-276 <ARZ>
A:Cross-references: EMBL:X75782; NID:g414736; PIDN:CAA53426.1; PID:g414737
A:Note: it is uncertain whether Met-1, Met-14 or Met-26 is the initiator
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke
submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome II BAC F26C24 genomic sequence.
A:Reference number: Z14680
A:Accession: T02601
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-276 <ROU>
A:Cross-references: EMBL:AC004705; NID:g3252804; PIDN:AAC24182.1; PID:g3252812
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84521
A:Status: preliminary
A:Molecule type: DNA
```

A:Residues: 1-276 <STO>
A:Cross-references: GB:AB002093; NID:g3252812; PIDN:AAC24182.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g14750; F26C24.11
A:Map position: 2

A:Genome: nuclear
A:Introns: 66/1; 113/1; 163/1; 202/3; 231/1; 246/3
C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology
C:Keywords: ATP; P-loop; phosphotransferase; purine nucleotide binding
F:1-37/Domain: transit peptide (chloroplast) #status predicted <TNP>
F:38-276/Product: adenylylsulfate kinase #status predicted <KAT>
F:101-262/Domain: adenylylsulfate kinase homology <ASK>
F:108-115/Region: nucleotide-binding motif A (P-loop)

Query Match 2.0%; Score 12; DB 1; Length 276;
Best Local Similarity 100.0%; Pred. No. 0.00065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LDGDNVRHGLNR 87

|||||

Db 135 LDGDNVRHGLNR 146

RESULT 15

T06100

adenylyl-sulfate kinase (EC 2.7.1.25) [validated] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 03-Jun-2002

C:Accession: T06100; T52055

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15184

A:Accession: T06100

A:Molecule type: DNA

A:Residues: 1-293 <BEV>

A:Cross-references: EMBL:AL035708

A:Experimental source: cultivar Columbia; BAC clone T5J17

R:Arz, H.E.; Gisselmann, G.; Schiffmann, S.; Schwenk, J.D.

Biochim. Biophys. Acta 1218, 447-452, 1994

A:Title: A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana.

A:Reference number: S47640; MUID:94325358; PMID:8049272

A:Accession: T52055

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-293 <ARZ>

A:Cross-references: EMBL:AF043351; PIDN:AAC39520.1

C:Genetics:

A:Gene: akn2

A:Map position: 4

A:Introns: 78/1; 177/1; 216/3; 245/1; 260/3

A:Note: T5J17.110

C:Function:

A:Description: EC 2.7.1.25 [validated, MUID:94325358]

C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology

C:Keywords: ATP; P-loop; phosphotransferase; purine nucleotide binding

F:115-279/Domain: adenylylsulfate kinase homology <ASK>

Query Match 2.0%; Score 12; DB 2; Length 293;

Best Local Similarity 100.0%; Pred. No. 0.00069;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 LDGDNVRHGLNR 87

|||||

Db 149 LDGDNVRHGLNR 160

Search completed: May 30, 2003, 12:33:08

Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:21:10 ; Search time 24 Seconds
(without alignments)
1062.831 Million cell updates/sec

Title: US-09-898-165B-7

Perfect score: 615

Sequence: 1 MSGIKKQKTEKQKSTNVVY.....MAPKAWKVLTDYRSEMDKN 615

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 10

Total number of hits satisfying chosen parameters: 21

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	509	82.8	614	1	PPS2_HUMAN
2	81	13.2	621	1	PPS2_MOUSE
3	38	6.2	624	1	PPS1_CAVPO
4	38	6.2	624	1	PPS1_HUMAN
5	38	6.2	624	1	PPS1_MOUSE
6	15	2.4	206	1	KAPS_EMENI
7	15	2.4	211	1	KAPS_PENCH
8	13	2.1	177	1	CYSC_SYNY3
9	12	2.0	276	1	KAP1_ARATH
10	12	2.0	293	1	KAP2_ARATH
11	12	2.0	623	1	CYSN_XYLFA
12	12	2.0	632	1	NODQ_RHITR
13	12	2.0	633	1	NODQ_RHISB
14	12	2.0	641	1	NODQ_RHIME
15	12	2.0	646	1	NODQ_RHIS3
16	11	1.8	110	1	YFMT_THETH
17	11	1.8	196	1	CYCL_PSEAE
18	11	1.8	214	1	CYC2_PSEAE
19	11	1.8	312	1	KAPS_CATRO
20	10	1.6	200	1	CYSC_ECOLI
21	10	1.6	215	1	CYSC_VIBCH
21	10	1.6	215	1	Q9kp21 vibrio chol

ALIGNMENTS

RESULT 1
PPS2_HUMAN
ID PPS2_HUMAN STANDARD; 614 AA.
AC Q95340; Q9UP30; Q9BZL2; Q9P0G6; Q9UHM1; Q9UKD3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 2 (PAPS
DE synthetase 2) (PAPSS 2) (Sulfurylase kinase 2) (SK2) (SK 2)
DE [Includes: Sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylyl
DE transferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase
DE (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)
DE (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-
DE phosphoadenosine-5'-phosphosulfate synthetase)].
GN PAPSS2 OR ATPSK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Fetal cartilage;
RX MEDLINE=98442651; PubMed=9771708;
RA Ul Haque M.F., King L.M., Krakow D., Cantor R.M., Rusiniak M.E.,
RA Swank R.T., Superti-Furga A., Haque S., Abbas H., Ahmad W., Ahmad M.,
RA Cohn D.H.;
RT "Mutations in orthologous genes in human spondyloepimetaphyseal
RT dysplasia and the brachymorphic mouse.";
RL Nat. Genet. 20:157-162(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A).
RA Franzon V.L., Gibson M.A., Hatzinikolas G., Cleary E.G., Woolatt E.,
RA Sutherland G.R.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM A).
RA Fuda H., Shimizu C., Strott C.A.;
RT "Human bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase:
RT differential expression of isoforms and effect of polymorphisms on
RT activity.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=20145452; PubMed=10679223;
RA Xu Z.-H., Otterness D.M., Freimuth R.R., Carlini E.J., Wood T.C.,
RA Mitchell S., Moon E., Kim U.-J., Xu J.-P., Siciliano M.J.,
RA Weinshilboum R.M.;
RT "Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSS1) and
RT PAPSS2: gene cloning, characterization and chromosomal localization.";
RL Biochem. Biophys. Res. Commun. 268:437-444(2000).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM B).
RC TISSUE=Liver;
RX Kurima K., Singh B., Schwartz N.B.;
RT "Genomic organization of the mouse and human genes encoding the ATP
RT sulfurylase/adenosine 5'-phosphosulfate kinase isoform SK2.";
RL J. Biol. Chem. 274:33306-33312(1999).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM B).
RA Venkatachalam K.V., Fuda H., Strott C.A.;
RT "3'-phosphoadenosine 5'-phosphosulfate synthase 2b isoform.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS
CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE
CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE
CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE
CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS
CC YIELDING 3'-PHOSPHOADENYLSULFATE (PAPS: ACTIVATED SULFATE DONOR
CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF
CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-
CC ACTIVATION PATHWAY. MAY HAVE AN IMPORTANT ROLE IN SKELETOGENESIS
CC DURING POSTNATAL GROWTH (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-

phosphoadenylylsulfate.
 -!- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.
 -!- ALTERNATIVE PRODUCTS: 2 isoforms; a (shown here) and b; are produced by alternative splicing.
 -!- TISSUE SPECIFICITY: EXPRESSED IN CARTILAGE.
 -!- DISEASE: DEFECTS IN PAPSS2 ARE THE CAUSE OF SPONDYLOEPIMETAPHYSEAL DYSPLASIA, PAKISTANI TYPE (SEMD), AN AUTOSOMAL RECESSIVE DISEASE. THIS FORM OF SEMD IS CHARACTERIZED BY A DYSPLASIA THAT IS PRIMARILY EPIPHYSEAL WITH ONLY MILD METAPHYSEAL ABNORMALITIES.
 -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE FAMILY.
 -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE ADENYLYLTRANSFERASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; AF091242; AAC64583.1; -
 EMBL; AF074331; AAD38423.1; -
 EMBL; AF313907; AAK00296.1; -
 EMBL; AF160509; AAF40307.2; -
 EMBL; AF160503; AAF40307.2; JOINED.
 EMBL; AF160504; AAF40307.2; JOINED.
 EMBL; AF160505; AAF40307.2; JOINED.
 EMBL; AF160506; AAF40307.2; JOINED.
 EMBL; AF160507; AAF40307.2; JOINED.
 EMBL; AF160508; AAF40307.2; JOINED.
 EMBL; AF173365; AAF12761.1; -
 EMBL; AF150754; AAF20366.2; -
 EMBL; BC009894; AAH09894.1; -
 Genes; HGNC:8604; PAPSS2.
 MIM; 603003; -
 InterPro; IPR002891; APS_kinase.
 InterPro; IPR002650; ATP-sulfurylase.
 Pfam; PF01583; APS_kinase; 1.
 Pfam; PF01747; ATP-sulfurylase; 1.
 ProDom; PD002350; APS_kinase; 1.
 ProDom; PD002381; ATP-sulfurylase; 1.
 TRAFs; TRAF00455; apsk; 1.
 Transferase; Nucleotidyltransferase; Kinase; Multifunctional enzyme; ATP-binding; Multigene family; Alternative splicing.
 KW ADENYLYLSULFATE KINASE.
 FT DOMAIN 1 2210
 FT NP_BIND 49 56
 FT ACT_SITE 123 123
 FT SITE 511 515
 FT VARSPLIC 288 288
 FT CONFLICT 166 166
 FT CONFLICT 361 361
 FT CONFLICT 426 426
 FT CONFLICT 567 567
 SQ SEQUENCE 614 AA; 69500 MW; 52F4B6D972DDA91E CRC64;
 Query Match 82.8%; Score 509; DB 1; Length 614;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 609; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSGIKKQKTENQKSTNNVYQAHVSRNKRGGVGTGRCGTGTCVWLTGLSGAGKTTISF 60
 DB 1 MSGIKKQKTENQKSTNNVYQAHVSRNKRGGVGTGRCGTGTCVWLTGLSGAGKTTISF 60
 QY 61 ALEEYLVSIAICYSILDGDNVRHGLNRNLNFGSPGDRNIRIAEVAKLFDAGLVCITS 120
 DB 61 ALEEYLVSIAICYSILDGDNVRHGLNRNLNFGSPGDRNIRIAEVAKLFDAGLVCITS 120

QY 121 FTSPFAKDRENARKTHESAGLFFFEIFVDAPLNICESRDVKGLYKRRAGEIKGFTGIDS 180
 DB 121 FTSPFAKDRENARKTHESAGLFFFEIFVDAPLNICESRDVKGLYKRRAGEIKGFTGIDS 180
 QY 181 DYEKETPPELVKTNLSTVSDCVHQVVELLQONIVPYTIIKDIHELFPENKLDHVAE 240
 DB 181 DYEKETPPELVKTNLSTVSDCVHQVVELLQONIVPYTIIKDIHELFPENKLDHVAE 240
 QY 241 AETLPSLSITKLDLQWQVLSGEGWATPLKGFMRKEKYLQVMHFDTLLDDGVNMSPIVL 300
 DB 241 AETLPSLSITKLDLQWQVLSGEGWATPLKGFMRKEKYLQVMHFDTLLDDGVNMSPIVL 300
 QY 301 PVSAEDKTRLEGCSKFVLAHGGRRVAILRDAEFYEHKREKRSRVWGTTCTKHPIKMYM 360
 DB 301 PVSAEDKTRLEGCSKFVLAHGGRRVAILRDAEFYEHKREKRSRVWGTTCTKHPIKMYM 360
 QY 361 ESGDWLVGGDLQVLEKIRWNGDGLQYRLTPLELKQCKEMNADAVFAFOLRNPVHNGHAL 420
 DB 361 ESGDWLVGGDLQVLEKIRWNGDGLQYRLTPLELKQCKEMNADAVFAFOLRNPVHNGHAL 420
 QY 421 LMQDTCRRLLERGGYKHPVLLHPLGQWTKDDVDPLDWRMKQHAALVEEGVLDPKSTIVAI 480
 DB 421 LMQDTCRRLLERGGYKHPVLLHPLGQWTKDDVDPLDWRMKQHAALVEEGVLDPKSTIVAI 480
 QY 481 FPPSMLYAGPTEVQVHCRSRMITAGANFYIVGRDPAGMPHPETKKDLYEPHGGKVLSPAP 540
 DB 481 FPPSMLYAGPTEVQVHCRSRMITAGANFYIVGRDPAGMPHPETKKDLYEPHGGKVLSPAP 540
 QY 541 GLTSEIIPFVAAYNKAKKAMDFDPARHNEFDFTSGTRMKRLAREGENPDPGFWAPKA 600
 DB 541 GLTSEIIPFVAAYNKAKKAMDFDPARHNEFDFTSGTRMKRLAREGENPDPGFWAPKA 600
 QY 601 WKVLTDYYS 610
 DB 601 WKVLTDYYS 610
 RESULT 2
 PPS2_MOUSE STANDARD; PRT; 621 AA.
 AC O88428; Q92274;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 2 (PAPS synthetase 2) (PAPSS 2) (Sulfurylase kinase 2) (SK2) (SK 2) [includes: Sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylyltransferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase) (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-phosphoadenosine-5'-phosphosulfate synthetase)].
 GN Mus musculus (Mouse).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT BM ARG-79.
 RC TISSUE=Liver;
 RX MEDLINE=98337975; PubMed=9671738;
 RA Kurima K., Warman M.L., Krishnan S., Domowicz M., Krueger R.C. Jr., Deyrup A., Schwartz N.B.;
 RT "A member of a new family of sulfate activating enzymes causes murine brachymorphism."; Proc. Natl. Acad. Sci. U.S.A. 95:8681-8685(1998).
 RL [2]
 RN SEQUENCE FROM N.A., VARIANT BM ARG-79, AND VARIANT LYS-109.
 RP STRAIN=PWK; TISSUE=Spleen;
 RX MEDLINE=98442651; PubMed=9771708;
 RA Ul Haque M.F., King L.M., Cantor R.M., Rusiniak M.E., Swank R.T., Superti-Furga A., Haque S., Abbas H., Ahmad W., Ahmad M., Cohn D.H.;
 RT "Mutations in orthologous genes in human spondyloepimetaphyseal dysplasia and the brachymorphic mouse.";

RL Nat. Genet. 20:157-162(1998).

CC -!- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS

CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE

CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE

CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE

CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS

CC YIELDING 3'-PHOSPHOADENYLSULFATE (PAPS): ACTIVATED SULFATE DONOR

CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF

CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-

CC ACTIVATION PATHWAY. MAY HAVE A IMPORTANT ROLE IN SKELETOGENESIS

CC DURING POSTNATAL GROWTH.

CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.

CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-

CC phosphoadenylylsulfate.

CC -!- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO

CC THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION

CC OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.

CC -!- TISSUE SPECIFICITY: EXPRESSED IN LIVER, CARTILAGE, SKIN AND BRAIN.

CC -!- DISEASE: DEFECTS IN PAPS2 ARE THE CAUSE OF BRACHYMRPHISM (BM), A

CC AUTOSOMAL RECESSIVE DISEASE, WHICH IS CHARACTERIZED BY ABNORMAL

CC HEPATIC DETOXIFICATION, BLEEDING TIMES AND POSTNATAL GROWTH, SUCH

CC AS DOME-SHAPED SKULL, SHORT THICK TAIL, AND SHORTENED BUT NOT

CC WIDENED LIMBS. THE ABNORMAL POSTNATAL GROWTH HAS BEEN ATTRIBUTED

CC TO UNDERSULFATION OF CARTILAGE PROTEOGLYCANS.

CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE

CC FAMILY.

CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE

CC ADENYLYLSULFATASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation.

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF052453; AAC40191.1; -

DR EMBL; AF085144; AAC98687.1; -

DR MGD; MGI:1330223; Paps2.

DR InterPro; IPR002891; APS_Kinase.

DR Pfam; PF01583; APS_kinase; 1

DR Pfam; PF01747; ATP-sulfurylase; 1

DR ProDom; PD002350; APS_Kinase; 1

DR ProDom; PD002381; ATP-sulfurylase; 1

DR TIGRFAMs; TIGR00455; apsk; 1

DR ATP-binding; Multigene family; Disease mutation; Polymorphism.

DR ADENYLYLSULFATE KINASE.

FT DOMAIN 1 7211

FT DOMAIN 7212 621

FT NP_BIND 50 57

FT ACT_SITE 124 124

FT SITE 517 521

FT VARIANT 79 79

FT VARIANT 109 109

FT CONFLICT 5 5

FT CONFLICT 290 294

FT CONFLICT 621 AA; 70290 MW; DE375F06A79FEAB CRC64;

FT SEQUENCE 621 AA; 70290 MW; DE375F06A79FEAB CRC64;

Query Match 13.2%; Score 81; DB 1; Length 621;

Best Local Similarity 100.0%; Pred. No. 5.4e-75;

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 LFADAGLVCTISFTSPAKDRENARKTHESAGLPFFEIFVDAPLNICESRDVKGLYKRR 168

DB 110 LFADAGLVCTISFTSPAKDRENARKTHESAGLPFFEIFVDAPLNICESRDVKGLYKRR 169

QY 169 AGEIKGFTGIDSDYKPTETPE 189

DB 170 AGEIKGFTGIDSDYKPTETPE 190

RESULT 3

PPSI_CAVPO STANDARD; PRT; 624 AA.

AC 054820;

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPS

DE synthetase 1) (PAPS1) (Sulfurylase kinase 1) (SK1) (SK 1)

DE [Includes: sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylylate

DE transferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase

DE (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)

DE (Adenosine 5'-phosphosulfate 3'-phosphotransferase) (3'-

DE phosphoadenosine-5'-phosphosulfate synthetase)]

GN PAPS1 OR PAPS OR ATPSKI.

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.

OX NCBI_TaxID:10141;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-NIH 2;

RA Venkatchalam K.V., Akita H., Strott C.A.;

RL Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.

CC -!- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS

CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE

CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE

CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE

CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS

CC YIELDING 3'-PHOSPHOADENYLSULFATE (PAPS): ACTIVATED SULFATE DONOR

CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF

CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-

CC ACTIVATION PATHWAY. ALSO INVOLVED IN THE BIOSYNTHESIS OF SULFATED

CC L-SELECTIN LIGANDS IN ENDOTHELIAL CELLS (BY SIMILARITY).

CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.

CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-

CC phosphoadenylylsulfate.

CC -!- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO

CC THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION

CC OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.

CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE

CC FAMILY.

CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE

CC ADENYLYLSULFATASE FAMILY.

CC -----

CC This SWISS-PROT entry is Copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation.

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AF004875; AAC02266.1; -

DR InterPro; IPR002891; APS_Kinase.

DR InterPro; IPR002650; ATP-sulfurylase.

DR Pfam; PF01583; APS_kinase; 1

DR Pfam; PF01747; ATP-sulfurylase; 1

DR ProDom; PD002350; APS_Kinase; 1

DR ProDom; PD002381; ATP-sulfurylase; 1

DR TIGRFAMs; TIGR00455; apsk; 1

DR Transferase; Nucleotidyltransferase; Kinase; Multifunctional enzyme;

DR ATP-binding; Multigene family.

FT DOMAIN 1 7220

FT DOMAIN 7221 624

FT NP_BIND 59 66

FT ACT_SITE 133 133

FT SITE 521 525

FT SEQUENCE 624 AA; 70395 MW; BF7461B4D07E2131 CRC64;

Query Match 6.2%; Score 38; DB 1; Length 624;

```
Best Local Similarity 100.0%; Pred. No. 1e-30; Mismatches 0; Indels 0; Gaps 0;
Matches 38; Conservative 0;

QY 20 YQAHVSRNKRQGVCTRGFRGCTVWLTGLSGAGKTT 57
      |||||
DB 30 YQAHVSRNKRQGVCTRGFRGCTVWLTGLSGAGKTT 67

RESULT 4
PPSI_HUMAN STANDARD; PRT; 624 AA.
ID 043252; 043841; 075332;
AC 30-MAY-2000 (Rel. 39, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPS
DE synthetase 1) (PAPSS 1) (Sulfurylase kinase 1) (SK1) (SK 1)
DE [Includes: Sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylyl-
DE transferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase
DE (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)
DE (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-
DE phosphoadenosine-5'-phosphosulfate synthetase)].
DE PAPSS1 OR PAPSS OR ATPSK1.
GN Homo sapiens (human)
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tonsil;
RX MEDLINE=98336023; PubMed=9576487;
RA Girard J.-P., Baekevold E.S., Amalric F.;
RT "Sulfation in high endothelial venules: cloning and expression of the
RT human PAPS synthetase."
RL FASEB J. 12:603-612(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=98334672; PubMed=9668121;
RA Venkatachalam K.V., Akita H., Strott C.A.;
RT "Molecular cloning, expression, and characterization of human
RT bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase and its
RT functional domains."
RL J. Biol. Chem. 273:19311-19320(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98312048; PubMed=9648242;
RA Yanagisawa K., Sakakibara Y., Suiko M., Takami Y., Nakayama T.,
RA Nakajima H., Takayanagi K., Natori Y., Liu M.-C.;
RT "cDNA cloning, expression, and characterization of the human
RT bifunctional ATP sulfurylase/adenosine 5'-phosphosulfate kinase
RT enzyme."
RL Biosci. Biotechnol. Biochem. 62:1037-1040(1998).
RN [4]
RP MUTAGENESIS OF HIS-151; ASN-426; GLY-427 AND HIS-428.
RX MEDLINE=99115594; PubMed=9915785;
RA Venkatachalam K.V., Fuda H., Koonin E.V., Strott C.A.;
RT "Site-selected mutagenesis of a conserved nucleotide binding HXGH
RT motif located in the ATP sulfurylase domain of human bifunctional
RT 3'-phosphoadenosine 5'-phosphosulfate synthase."
RL J. Biol. Chem. 274:2601-2604(1999).
RN [5]
RP FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS
CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE
CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE
CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE
CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS
CC YIELDING 3'-PHOSPHOADENYLYLSULFATE (PAPS: ACTIVATED SULFATE DONOR
CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF
CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-
CC ACTIVATION PATHWAY. ALSO INVOLVED IN THE BIOSYNTHESIS OF SULFATED
CC L-SELECTIN LIGANDS IN ENDOTHELIAL CELLS.
CC -!- CATALYTIC ACTIVITY: ATP + sulfate -> diphosphate + adenylylsulfate.
CC
```

```
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate -> ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- ENZYME REGULATION: INHIBITED BY CHLORATE.
CC -!- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO
CC THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION
CC OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOLOGICS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN TESTIS, PANCREAS, KIDNEY, THYMUS,
CC PROSTATE, OVARY, SMALL INTESTINE, COLON, LEUKOCYTES AND LIVER.
CC ALSO EXPRESSED IN HIGH ENDOTHELIAL VENULES (HEV) CELLS AND IN
CC CARTILAGE.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE
CC FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE
CC ADENYLYLTRANSFERASE FAMILY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Y10387; CAA71413.1;
CC EMBL: U53447; AAC39894.1;
CC EMBL: AF033026; AAC28429.1;
CC Genbank: HGNC:8603; PAPSS1.
CC MIM: 603262;
CC
CC InterPro: IPR002891; APS_kinase.
CC InterPro: IPR002650; ATP-sulfurylase.
CC Pfam: PF01583; APS_kinase; 1.
CC Pfam: PF01747; ATP-sulfurylase; 1.
CC ProDom: PD002350; APS_kinase; 1.
CC ProDom: PD002381; ATP-sulfurylase; 1.
CC TIGRFAMs: TIGR00455; apsk; 1.
CC Transfaser: Nucleotidyltransferase; Kinase; Multifunctional enzyme;
CC ATP-binding; Multigene family.
CC DOMAIN 1 7220 ADENYLYLSULFATE KINASE.
CC FT DOMAIN 7221 624 SULFATE ADENYLYLTRANSFERASE.
CC FT NP_BIND 59 66 ATP (POTENTIAL).
CC FT ACT_SITE 133 133 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
CC SIMILARITY).
CC FT SITE 521 525 PP-MOTIF (BY SIMILARITY).
CC FT MUTAGEN 425 425 H->A: LOSS OF ACTIVITY.
CC FT MUTAGEN 426 426 N->K: INCREASED ACTIVITY.
CC FT MUTAGEN 427 427 G->A: 30% DECREASE IN ACTIVITY.
CC FT MUTAGEN 428 428 H->A: LOSS OF ACTIVITY.
CC FT MUTAGEN 427 428 GH->AA: LOSS OF ACTIVITY.
CC FT CONFLICT 270 270 L -> F (IN REF. 2).
CC FT CONFLICT 456 456 MISSING (IN REF. 2 AND 3).
CC FT CONFLICT 587 587 S -> L (IN REF. 2).
CC SQ SEQUENCE 624 AA; 70847 MW; 6BC4F9648016CA31 CRC64;

Query Match 6.2%; Score 38; DB 1; Length 624;
Best Local Similarity 100.0%; Pred. No. 1e-30;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRQGVCTRGFRGCTVWLTGLSGAGKTT 57
      |||||
DB 30 YQAHVSRNKRQGVCTRGFRGCTVWLTGLSGAGKTT 67

RESULT 5
PPSI_MOUSE STANDARD; PRT; 624 AA.
ID PPS1_MOUSE
AC Q60967;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPS
DE synthetase 1) (PAPSS 1) (Sulfurylase kinase 1) (SK1) (SK 1)
DE [Includes: Sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylyl-
DE transferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase
DE (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)
DE (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-
DE phosphoadenosine-5'-phosphosulfate synthetase)].
DE PAPSS1 OR PAPSS OR ATPSK1.
GN Mus musculus (house mouse)
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tonsil;
RX MEDLINE=98336023; PubMed=9576487;
RA Girard J.-P., Baekevold E.S., Amalric F.;
RT "Sulfation in high endothelial venules: cloning and expression of the
RT human PAPS synthetase."
RL FASEB J. 12:603-612(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=98334672; PubMed=9668121;
RA Venkatachalam K.V., Akita H., Strott C.A.;
RT "Molecular cloning, expression, and characterization of human
RT bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase and its
RT functional domains."
RL J. Biol. Chem. 273:19311-19320(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98312048; PubMed=9648242;
RA Yanagisawa K., Sakakibara Y., Suiko M., Takami Y., Nakayama T.,
RA Nakajima H., Takayanagi K., Natori Y., Liu M.-C.;
RT "cDNA cloning, expression, and characterization of the human
RT bifunctional ATP sulfurylase/adenosine 5'-phosphosulfate kinase
RT enzyme."
RL Biosci. Biotechnol. Biochem. 62:1037-1040(1998).
RN [4]
RP MUTAGENESIS OF HIS-151; ASN-426; GLY-427 AND HIS-428.
RX MEDLINE=99115594; PubMed=9915785;
RA Venkatachalam K.V., Fuda H., Koonin E.V., Strott C.A.;
RT "Site-selected mutagenesis of a conserved nucleotide binding HXGH
RT motif located in the ATP sulfurylase domain of human bifunctional
RT 3'-phosphoadenosine 5'-phosphosulfate synthase."
RL J. Biol. Chem. 274:2601-2604(1999).
RN [5]
RP FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS
CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE
CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE
CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE
CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS
CC YIELDING 3'-PHOSPHOADENYLYLSULFATE (PAPS: ACTIVATED SULFATE DONOR
CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF
CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-
CC ACTIVATION PATHWAY. ALSO INVOLVED IN THE BIOSYNTHESIS OF SULFATED
CC L-SELECTIN LIGANDS IN ENDOTHELIAL CELLS.
CC -!- CATALYTIC ACTIVITY: ATP + sulfate -> diphosphate + adenylylsulfate.
CC
```

DE (EC 2.7.1.25) (adenylylsulfate 3'-phosphotransferase) (APS kinase)
DE (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-
DE phosphoadenosine-5'-phosphosulfate synthetase)]].
GN PAPSS1 OR PAPSS OR ASAPK OR ATPSK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RC MEDLINE=96094345; PubMed=7493984;
RA Li H., Deyrup A., Mensch J.R. Jr., Domowicz M., Konstantinidis A.K.,
RA Schwartz N.B.;
RT "The isolation and characterization of cDNA encoding the mouse
RT bifunctional ATP sulfolysase-adenosine 5'-phosphosulfate kinase.";
RL J. Biol. Chem. 270:29453-29459(1995).
CC -!- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS
CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE
CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE
CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE
CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS
CC YIELDING 3'-PHOSPHOADENYLYLSULFATE (PAPS). ACTIVATED SULFATE DONOR
CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF
CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-
CC ACTIVATION PATHWAY. ALSO INVOLVED IN THE BIOSYNTHESIS OF SULFATED
CC L-SELECTIN LIGANDS IN ENDOTHELIAL CELLS.
CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO
CC THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION
CC OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE NEONATAL BRAIN AND IN
CC CARTILAGE.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE
CC FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE
CC ADENYLYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U34883; AAC52328.1; -
DR MGI: 1330587; Papsl1.
DR InterPro: IPR002891; APS_kinase.
DR Pfam: PF01583; APS_kinase; 1.
DR Pfam: PF01747; ATP-sulfurylase; 1.
DR ProDom: PD002350; APS_kinase; 1.
DR ProDom: PD002381; ATP-sulfurylase; 1.
DR TIGRFAMs: TIGR00455; apsk; 1.
DR Transferase; Nucleotidyltransferase; Kinase; Multifunctional enzyme;
DR ATP-binding; Multigene family.
DR DOMAIN 1 2220 ADENYLYLSULFATE KINASE.
DR DOMAIN 2221 624 SULFATE ADENYLYLTRANSFERASE.
DR NP_BIND 59 66 ATP (POTENTIAL).
DR ACT_SITE 133 133 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
DR PP-MOTIF.
DR SITE 521 525 PP-MOTIF.
DR SEQUENCE 624 AA; 70794 MW; B487EFAF9B78BE3E CRC64;
Query Match 6.28; Score 38; DB 1; Length 624;
Best Local Similarity 100.0%; Pred. No. 1e-30;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 YQAHVSRNKRQGVVTRGFRGCTVWLTGLSGAGKTT 57
|||||

Db 30 YQAHVSRNKRQGVVTRGFRGCTVWLTGLSGAGKTT 67
RESULT 6
KAPS_EMENI STANDARD; PRT; 206 AA.
AC Q92203;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
DE 5'-phosphosulfate kinase) (ATP adenosine 5'-phosphosulfate 3'-
DE phosphotransferase).
GN SD.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RA Clarke D.L., Newbert R.W., Turner G.;
RT "Cloning and characterisation of the APS kinase gene from
RT Aspergillus nidulans.";
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ASSIMILATION PATHWAY THAT
CC LEADS TO THE BIOSYNTHESIS OF METHIONINE AND CYSTEINE.
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y08866; CAA70089.1; -
DR InterPro: IPR002891; APS_Kinase.
DR Pfam: PF01583; APS_Kinase; 1.
DR ProDom: PD002350; APS_Kinase; 1.
DR TIGRFAMs: TIGR00455; apsk; 1.
DR Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
DR NP_BIND 31 38 ATP (POTENTIAL).
DR ACT_SITE 105 105 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
DR SIMILARITY).
DR SEQUENCE 206 AA; 23028 MW; 6AADD483E2BCA1CD CRC64;
Query Match 2.48; Score 15; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 98 ENIRRIAEVAKLFAD 112
|||||
Db 80 ENIRRIAEVAKLFAD 94
RESULT 7
KAPS_PENCH STANDARD; PRT; 211 AA.
AC Q12657;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
DE 5'-phosphosulfate kinase) (ATP adenosine 5'-phosphosulfate 3'-
DE phosphotransferase).
OS Penicillium chrysogenum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.

```

OX NCBI_TaxID=5076;
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 24791;
RA Foster B.A.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-phosphoadenylylsulfate.
CC -1- PATHWAY: SECOND STEP IN THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO THE BIOSYNTHESIS OF METHIONINE AND CYSTEINE.
CC -1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39393; AAA81521.1;
DR InterPro; IPR002891; APS_Kinase.
DR Pfam; PF01583; APS_Kinase; 1.
DR ProDom; PD002350; APS_Kinase; 1.
DR TIGRFAMs; TIGR00455; apsk; 1.
KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation.
FT NP_BIND 32 39 ATP (POTENTIAL).
FT ACT_SITE 107 107 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY SIMILARITY).
FT SEQUENCE 211 AA; 23770 MW; 7DDC4BDA867FE7C2 CRC64;

Query Match 2.48; Score 15; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 ENIRRIAEVAKLFAD 112
DB 82 ENIRRIAEVAKLFAD 96
|||||
|||||

RESULT 8
CYSC SYN3
ID CYSC SYN3 STANDARD; PRT; 177 AA.
AC P72940;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Probable adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase).
DE CYSC OR SLR0676.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hiroseawa M., Sugiyura M., Sasamoto S., Kimura T., Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-phosphoadenylylsulfate.
CC -1- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE

```

```

CC BIOSYNTHETIC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90902; BAA16957.1;
DR InterPro; IPR002891; APS_Kinase.
DR Pfam; PF01583; APS_Kinase; 1.
DR ProDom; PD002350; APS_Kinase; 1.
DR TIGRFAMs; TIGR00455; apsk; 1.
KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation; Complete proteome.
FT NP_BIND 12 19 ATP (BY SIMILARITY).
FT ACT_SITE 86 86 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY SIMILARITY).
FT SEQUENCE 177 AA; 19674 MW; E24D5B5452FCD7D8 CRC64;

Query Match 2.18; Score 13; DB 1; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 WLTLGSLGAGKTTI 58
DB 9 WLTLGSLGAGKTTI 21
|||||
|||||

RESULT 9
KAP1 ARATH
ID KAP1 ARATH STANDARD; PRT; 276 AA.
AC Q43295;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenylylsulfate kinase 1, chloroplast precursor (EC 2.7.1.25) (APS kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase).
DE AKN1 OR AT2G14750 OR F26C24.11 OR T6B13.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Leaf;
RX MEDLINE=943325358; PubMed=8049272;
RA Arz H.E., Gisselmann G., Schiffmann S., Schwenk J.-D.;
RT "A cDNA for adenylyl sulphate (APS)-kinase from Arabidopsis thaliana.";
RL Biochim. Biophys. Acta 1218:447-452(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94345022; PubMed=8066145;
RA Jain A., Leustek T.;
RT "A cDNA clone for 5'-adenylylphosphosulfate kinase from Arabidopsis thaliana.";
RL Plant Physiol. 105:771-772(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta;
RA Leustek T.;
RT "Analysis of ATP-sulfurylase and APS-kinase genomic DNA sequences in Arabidopsis thaliana.";
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;

```

```

RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana";
RL Nature 402:761-768(1999).
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
CC BIOSYNTHETIC PATHWAY.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X75782; CAA53426.1; -
DR EMBL; U52338; AAC50035.1; -
DR EMBL; U59759; AAC50034.1; -
DR EMBL; AC004705; AAC24182.1; -
DR InterPro; IPR002891; APS_kinase.
DR Pfam; PF01583; APS_Kinase; 1.
DR ProDom; PD002350; APS_Kinase; 1.
DR TIGRFAMs; TIGR00455; apsk; 1.
DR Transphorase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation; Kinase; Transit peptide; Chloroplast.
FT TRANSIT 1 37
FT CHAIN 38 276
FT NP_BIND 108 115
FT ACT_SITE 182 182
FT FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 276 AA; 29787 MW; CB698643AA09D811 CRC64;
Query Match 2.0%; Score 12; DB 1; Length 276;
Best Local Similarity 100.0%; pred. No. 0.00029;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 76 LDGDNVRHGLNR 87
Db 135 LDGDNVRHGLNR 146
RESULT 10
KAP2_ARATH STANDARD; PRT; 293 AA.
AC 049196;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenylylsulfate kinase 2, chloroplast precursor (EC 2.7.1.25) (APS
DE kinase) (adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-
DE phosphosulfate 3'-phosphotransferase).
OS AKN2 OR AT4G39940 OR T517.110.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC STRAIN=cv. Columbia;
RA Schiffmann S., Schwenn J.-D.;
RT "Isolation of cDNA clones encoding adenosine-5'-phosphosulfate-kinase
RT (EC2.7.1.25) from Catharanthus roseus and an isoform (aku2) from
RT Arabidopsis.";
RL (In) Plant Gene Register PGR98-116.
RN [2]
RX SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansong W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,
RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohenseil J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
RA Weitzenecker T., Bothe G., Rampsperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Bernseiser S., Hempel S., Feldpausch M., Lambers S., Van den Daele H.,
RA De Keyser A., Buysshaert C., Giesen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlroy K., Mayes R.,
RA Pettett A., Rajadram M.A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehnert T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Grandrath K., Dauner D., Herzl A.,
RA Neumann S., Argirou A., Vitale D., Liguori R., Piravandi E.,
RA Massenel O., Quigley F., Clabaud G., Muendlein A., Felber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chefdr F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,
RA Perez-Perez A., Fumelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Hejlsman L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,
RA Freshman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Speth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antonoli B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vill D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana";
RL Nature 402:769-777(1999).
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
CC BIOSYNTHETIC PATHWAY.
CC -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

DR EMBL; AF043351; AAC39520.1; -
 DR EMBL; AL035708; CAB38907.1; -
 DR EMBL; AL161596; CAB80657.1; -
 DR InterPro; IPR002891; APS_Kinase.
 DR Pfam; PF01583; APS_Kinase; 1.
 DR ProDom; PD002350; APS_Kinase; 1.
 DR TIGRFAMs; TIGR00455; apsk; 1.
 KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
 FT Phosphorylation; Transient peptide; Chloroplast.
 FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
 FT CHAIN ? 293 ADENYLSULFATE KINASE 2.
 FT NP_BIND 142 149 ATP (POTENTIAL).
 FT ACT_SITE 196 196 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
 FT ACT_SITE 196 196 SIMILARITY).
 SQ SEQUENCE 293 AA; 31977 MW; 71FTAD6E9B026886 CRC64;
 Query Match 2.0%; Score 12; DB 1; Length 293;
 Best Local Similarity 100.0%; Pred. No. 0.0003;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 76 LDGDNVVRHGLNR 87
 D 149 LDGDNVVRHGLNR 160
 RESULT 11
 CYSN_XYLFA STANDARD; PRT; 623 AA.
 ID CYSN_XYLFA STANDARD; PRT; 623 AA.
 AC Q9PD78;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE CysN/cysc bifunctional enzyme [includes: Sulfate adenylyltransferase;
 DE subunit 1 (EC 2.7.7.4) (Sulfate adenylyltransferase) (SAT) (ATP-
 DE sulfurylase large subunit); Adenylylsulfate kinase (EC 2.7.1.25) (APS
 DE kinase) (ATP adenosine-5'-phosphosulfate 3'-phosphotransferase)].
 GN CYSNC OR XF1501.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; gamma subphylum; Xanthomonas group;
 OC Xylella.
 OC NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9a5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-159(2000).

CC -1- FUNCTION: ATP SULFURYLASE MAY BE THE GTPASE, REGULATING ATP
 CC SULFURYLASE ACTIVITY (BY SIMILARITY).
 CC -1- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE
 CC (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
 CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
 CC phosphoadenylylsulfate.
 CC -1- PATHWAY: FIRST AND SECOND STEPS IN THE SULFATE ACTIVATION PATHWAY.
 CC THESE REACTIONS OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE
 CC CYSTEINE BIOSYNTHETIC PATHWAY.
 CC -1- SUBUNIT: HETERODIMER COMPOSED OF CYSN, THE SMALLER SUBUNIT, AND
 CC CYSN (BY SIMILARITY).
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING
 CC ELONGATION FACTOR FAMILY. CYSN/NODOQ SUBFAMILY.
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE003980; AAF84310.1; ALT_INIT.
 CC HSSP; P02990; 1ETU.
 CC InterPro; IPR002891; APS_Kinase.
 CC InterPro; IPR004161; EFTU_D2.
 CC InterPro; IPR000795; EF_GTPbind.
 CC Pfam; PF00009; GTP_EFTU; 1.
 CC Pfam; PF01583; APS_Kinase; 1.
 CC Pfam; PF03144; GTP_EFTU_D2; 1.
 CC PRINTS; PR00315; ELONGATNFCT.
 CC ProDom; PD002350; APS_Kinase; 1.
 CC TIGRFAMs; TIGR00455; apsk; 1.
 CC PROSITE; PS00301; EFATOR_GTP; 1.
 KW Cysteine biosynthesis; Transferase; Nucleotidyltransferase;
 KW GTP-binding; Kinase; ATP-binding; Multifunctional enzyme;
 KW Complete proteome.
 KW DOMAIN 1 450 SULFATE ADENYLYL TRANSFERASE.
 FT NP_BIND 23 30 GTP (BY SIMILARITY).
 FT NP_BIND 102 106 GTP (BY SIMILARITY).
 FT NP_BIND 157 160 GTP (BY SIMILARITY).
 FT NP_BIND 459 466 ATP (POTENTIAL).
 FT ACT_SITE 533 533 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
 FT ACT_SITE 533 533 SIMILARITY).
 SQ SEQUENCE 623 AA; 68277 MW; C20730A365B28E94 CRC64;
 Query Match 2.0%; Score 12; DB 1; Length 623;
 Best Local Similarity 100.0%; Pred. No. 0.00058;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 76 LDGDNVVRHGLNR 87
 D 486 LDGDNVVRHGLNR 497
 RESULT 12
 NODOQ_RHTR STANDARD; PRT; 632 AA.
 ID NODOQ_RHTR STANDARD; PRT; 632 AA.
 AC P52978;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE NodQ bifunctional enzyme (Nodulation protein Q) [Includes: Sulfate
 DE adenylyltransferase subunit 1 (EC 2.7.7.4) (Sulfate adenylyl
 DE transferase) (SAT) (ATP-sulfurylase large subunit); Adenylylsulfate
 DE kinase (EC 2.7.1.25) (APS kinase) (ATP adenosine-5'-phosphosulfate 3'-
 DE phosphotransferase)].
 GN NODOQ.
 OS Rhizobium tropici.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=398;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-CFN 299;
 RC MEDLINE=96303335; PubMed=8755625;
 RA Laermans T., Caluwaerts I., Verreth C., Rogel M.A.,
 RT "Isolation and characterization of Rhizobium tropici Nod factor
 sulfation genes";
 RL Mol. Plant Microbe Interact. 9:492-500(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CIAT899;
 RX MEDLINE=97002746; PubMed=8850086;
 RA Folch-Mallol J.L., Marroqui S., Sousa C., Manyani H., Lopez-Lara I.M.,
 RA van der Drift K.M.G.M., Haverkamp J., Quinto C., Gil-Serrano A.,
 RA Thomas-Oates J., Spaink H.P., Megias M.;
 RT "Characterization of Rhizobium tropici CIAT899 nodulation factors: the
 role of nodH and nodP genes in their sulfation";
 RL Mol. Plant Microbe Interact. 9:151-163(1996).
 CC -!- FUNCTION: PROPOSED TO PROVIDE ACTIVATED SULFATE FOR TRANSFER TO
 CC NOD FACTOR. ATP SULFURYLASE MAY BE THE GTPASE, REGULATING ATP
 CC SULFURYLASE ACTIVITY (BY SIMILARITY).
 CC -!- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE
 CC (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
 CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
 CC phosphoadenylylsulfate.
 CC -!- SUBUNIT: SULFATE-ACTIVATING ENZYMES. NODP AND NODO, MAY BE
 CC PHYSICALLY ASSOCIATED.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING
 CC ELONGATION FACTOR FAMILY. CYSN/NODO SUBFAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U47272; AAB08984.1; -;
 DR EMBL; X87608; CAA60914.1; -;
 DR InterPro; IPR002891; APS_kinase.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF01583; APS_kinase; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR ProDom; PD002350; APS_kinase; 1.
 DR TIGRFAMs; TIGR00455; apsk; 1.
 DR PROSITE; PS00301; EFACITOR_GTP; 1.
 KW Nodulation; Transference; Nucleotidyltransferase; GTP-binding;
 KW Kinase; ATP-binding; Multifunctional enzyme.
 FT DOMAIN 1 457 SULFATE ADENYLYL TRANSFERASE.
 FT NP_BIND 458 632 ADENYLYLSULFATE KINASE.
 FT NP_BIND 31 38 GTP (BY SIMILARITY).
 FT NP_BIND 109 113 GTP (BY SIMILARITY).
 FT NP_BIND 164 167 GTP (BY SIMILARITY).
 FT NP_BIND 466 473 ATP (POTENTIAL).
 FT ACT_SITE 540 540 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
 FT SIMILARITY).
 SQ SEQUENCE 632 AA; 69814 MW; 8FA3DF0CC7C604B6 CRC64;
 Query Match 2.0%; Score 12; DB 1; Length 632;
 Best Local Similarity 100.0%; Pred. No. 0.00059;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 76 LDGDNVRHGLNR 87

Db 493 LDGDNVRHGLNR 504
 [1]
 RESULT 13
 NODO_RHISB
 ID NODO_RHISB STANDARD; PRT; 633 AA.
 AC 007309;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE NodQ bifunctional enzyme (Nodulation protein Q) [Includes: Sulfate
 DE adenylyltransferase subunit 1 (EC 2.7.7.4) (Sulfate adenylyl
 DE transferase) (SAT) (ATP-sulfurylase large subunit); Adenylylsulfate
 DE kinase (EC 2.7.1.25) (APS kinase) (ATP adenosine-5'-phosphosulfate 3'-
 DE phosphotransferase)].
 GN NODO.
 OS Rhizobium sp. (strain BR816).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=48291;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98083747; PubMed=9421916;
 RA Laermans T., Coolsaet N., Verreth C., Snoeck C., Hellings N.,
 RA Vanderleyden J., Martinez-Romero E.;
 RT "Functional redundancy of genes for sulphate activation enzymes in
 RT Rhizobium sp. BR816";
 RL Microbiology 143:3933-3942(1997).
 CC -!- FUNCTION: PROPOSED TO PROVIDE ACTIVATED SULFATE FOR TRANSFER TO
 CC NOD FACTOR. ATP SULFURYLASE MAY BE THE GTPASE, REGULATING ATP
 CC SULFURYLASE ACTIVITY (BY SIMILARITY).
 CC -!- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE
 CC (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
 CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
 CC phosphoadenylylsulfate.
 CC -!- SUBUNIT: SULFATE-ACTIVATING ENZYMES. NODP AND NODO, MAY BE
 CC PHYSICALLY ASSOCIATED (POTENTIAL).
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING
 CC ELONGATION FACTOR FAMILY. CYSN/NODO SUBFAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U59507; AAB95249.1; -;
 DR InterPro; IPR002891; APS_kinase.
 DR InterPro; IPR004161; EFTU_D2.
 DR InterPro; IPR000795; EF_GTPbind.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00009; GTP_EFTU; 1.
 DR Pfam; PF01583; APS_kinase; 1.
 DR Pfam; PF03144; GTP_EFTU_D2; 1.
 DR ProDom; PD002350; APS_kinase; 1.
 DR TIGRFAMs; TIGR00231; small_GTP; 1.
 DR TIGRFAMs; TIGR00455; apsk; 1.
 DR PROSITE; PS00301; EFACITOR_GTP; 1.
 KW Nodulation; Transference; Nucleotidyltransferase; GTP-binding;
 KW Kinase; ATP-binding; Multifunctional enzyme.
 FT DOMAIN 1 458 SULFATE ADENYLYL TRANSFERASE.
 FT NP_BIND 459 633 ADENYLYLSULFATE KINASE.
 FT NP_BIND 31 38 GTP (BY SIMILARITY).
 FT NP_BIND 110 114 GTP (BY SIMILARITY).
 FT NP_BIND 165 168 GTP (BY SIMILARITY).
 FT NP_BIND 467 474 ATP (POTENTIAL).
 FT ACT_SITE 541 541 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY

FT SQ SEQUENCE 633 AA; 69279 MW; 696BCA683D4B3111 CRC64;
SIMILARITY).
Query Match 2.0%; Score 12; DB 1; Length 633;
Best Local Similarity 100.0%; Pred. No. 0.00059;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 LDGDNVRHGLNR 87
|||||
DB 494 LDGDNVRHGLNR 505

RESULT 14
ID NODOQ_RHIME STANDARD; PRT; 641 AA.
AC P13442;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE NodQ bifunctional enzyme (Nodulation protein Q) [Includes: Sulfate
DE adenylyltransferase subunit 1 (EC 2.7.7.4) (Sulfate adenylyl-
DE transferase) (SAT) (ATP-sulfurylase large subunit); Adenylylsulfate
DE kinase (EC 2.7.1.25) (APS kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE phosphotransferase)].
DE NODOQ OR RA0469 OR SMA0857.
GN Rhizobium meliloti (Sinorhizobium meliloti).
OS Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / S047;
RX MEDLINE=89313304; PubMed=2546009;
RA Cervantes E., Sharma S.B., Mailliet F., Vasse J., Truchet G.,
RA Rosenberg C.;
RT "The Rhizobium meliloti host range nodQ gene encodes a protein which
RT shares homology with translation elongation and initiation factors.";
RL Mol. Microbiol. 3:745-755(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92288403; PubMed=2520820;
RA Schwedock J., Long S.R.;
RT "Nucleotide sequence and protein products of two new nodulation genes
RT of Rhizobium meliloti, nodP and nodQ.";
RL Mol. Plant Microbe Interact. 2:181-194(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9889(2001).
CC -1- FUNCTION: PROPOSED TO PROVIDE ACTIVATED SULFATE FOR TRANSFER TO
CC NOD FACTOR. ATP SULFURYLAASE MAY BE THE GTPASE, REGULATING ATP
CC SULFURYLAASE ACTIVITY (BY SIMILARITY).
CC -1- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -1- SUBUNIT: SULFATE-ACTIVATING ENZYMES, NODP AND NODOQ, MAY BE
CC PHYSICALLY ASSOCIATED (POTENTIAL).
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING
CC ELONGATION FACTOR FAMILY. CYSN/NODO SUBFAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE
CC FAMILY.
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14809; CAA32914.1; -;
DR EMBL; M68858; AAA26343.1; -;
DR EMBL; AE007237; AKK65127.1; -;
DR PIR; S14899; Z22RNO.
DR InterPro: IPR002891; APS_kinase.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR000795; EF_GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF01583; APS_kinase; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR ProDom; PD002350; APS_kinase; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR TIGRFAMs; TIGR00455; apsk; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Plasmid; Nodulation; Transferase; Nucleotidyltransferase; GTP-binding;
KW Kinase; ATP-binding; Multifunctional enzyme; Complete proteome.
FT DOMAIN 1 458 SULFATE ADENYLYL TRANSFERASE.
FT NP_BIND 31 38 ADENYLYLSULFATE KINASE.
FT NP_BIND 110 114 GTP (BY SIMILARITY).
FT NP_BIND 165 168 GTP (BY SIMILARITY).
FT NP_BIND 467 474 GTP (BY SIMILARITY).
FT ACT_SITE 524 524 ATP (POTENTIAL).
FT FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SIMILARITY).
SQ SEQUENCE 641 AA; 70614 MW; 1E1261F04ED33A93 CRC64;
Query Match 2.0%; Score 12; DB 1; Length 641;
Best Local Similarity 100.0%; Pred. No. 0.0006;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 76 LDGDNVRHGLNR 87
|||||
DB 494 LDGDNVRHGLNR 505

RESULT 15
ID NODOQ_RHIS3 STANDARD; PRT; 646 AA.
AC P72339;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NodQ bifunctional enzyme (Nodulation protein Q) [Includes: Sulfate
DE adenylyltransferase subunit 1 (EC 2.7.7.4) (Sulfate adenylyl-
DE transferase) (SAT) (ATP-sulfurylase large subunit); Adenylylsulfate
DE kinase (EC 2.7.1.25) (APS kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE phosphotransferase)].
GN NODOQ.
OS Rhizobium sp. (strain N33).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=103798;
RN [1]
RP SEQUENCE FROM N.A.
RL Cloutier J.;
RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROPOSED TO PROVIDE ACTIVATED SULFATE FOR TRANSFER TO
CC NOD FACTOR. ATP SULFURYLAASE MAY BE THE GTPASE, REGULATING ATP
CC SULFURYLAASE ACTIVITY (BY SIMILARITY).
CC -1- FUNCTION: APS KINASE CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC

CC -!- SUBUNIT: SULFATE-ACTIVATING ENZYMES, NODP AND NODQ, MAY BE
 CC PHYSICALLY ASSOCIATED (POTENTIAL).
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GTP-BINDING
 CC ELONGATION FACTOR FAMILY. CYSN/NODQ SUBFAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE APS KINASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics, and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U53327; AAB16902.1; -
 CC HSSP; P02990; 1ETU.
 CC InterPro; IPR002891; APS_Kinase.
 CC InterPro; IPR004161; EFTU_D2
 CC InterPro; IPR000795; EF_GTPbind.
 CC Pfam; PF00009; GTP_EFTU; 1.
 CC Pfam; PF01583; APS_Kinase; 1.
 CC Pfam; PF03144; GTP_EFTU_D2; 1.
 CC ProDom; PD002350; APS_Kinase; 1.
 CC TIGRFAMs; TIGR00455; apsk; 1.
 CC PROSITE; PS00301; EFACOR_GTP; 1.
 KW Nucleotidyltransferase; Nucleotidyltransferase; GTP-binding;
 KW Kinase; ATP-binding; Multifunctional enzyme.
 FT DOMAIN 1 461 SULFATE ADENYLYL TRANSFERASE.
 FT NP_BIND 34 41 ADENYLYLSULFATE KINASE.
 FT NP_BIND 113 117 GTP (BY SIMILARITY).
 FT NP_BIND 168 171 GTP (BY SIMILARITY).
 FT NP_BIND 470 477 ATP (POTENTIAL).
 FT ACT_SITE 544 544 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
 FT SIMILARITY).
 SQ SEQUENCE 646 AA; 71488 MW; 1DBEEIDA257FE128 CRC64;

Query Match 2.0%; Score 12; DB 1; Length 646;
 Best Local Similarity 100.0%; Pred. No. 0.0006;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 76 LDGDNVRHGLNR 87
 Db 497 LDGDNVRHGLNR 508
 |||||

Search completed: May 30, 2003, 12:30:43
 Job time : 24 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:26:25 ; Search time 87 Seconds
(without alignments)
1456.540 Million cell updates/sec

Title: US-09-898-165B-7
Perfect score: 615
Sequence: 1 MSGIKKQKTENOOKSTNVVY.....MAPKAWKVLTDYRSEMDKN 615

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 10

Total number of hits satisfying chosen parameters: 51

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertbrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142	23.1	265	4 Q9UIR2	Q9UIR2 homo sapien
2	81	13.2	621	11 Q9QYSO	Q9QYSO mus musculus
3	58	9.4	620	11 Q9TK86	Q9TK86 cavia porce
4	38	6.2	624	4 Q9UE98	Q9UE98 homo sapien
5	38	6.2	624	4 Q9E1P9	Q9E1P9 homo sapien
6	38	6.2	624	4 Q9GTF4	Q9GTF4 homo sapien
7	37	6.0	613	13 Q9QXY2	Q9QXY2 fugu rubrip
8	26	4.2	618	5 Q9NDP8	Q9NDP8 ciona intes
9	22	3.6	424	10 Q43170	Q43170 solanum tub
10	22	3.6	458	10 Q8W1X3	Q8W1X3 allium cepa
11	22	3.6	461	10 Q9SDE4	Q9SDE4 allium cepa
12	22	3.6	461	10 Q8SAG1	Q8SAG1 glycine max
13	22	3.6	463	10 Q42519	Q42519 arabidopsis
14	22	3.6	463	10 Q9SE02	Q9SE02 arabidopsis
15	22	3.6	463	10 Q43183	Q43183 solanum tub
16	22	3.6	463	10 Q9L1K9	Q9L1K9 arabidopsis

ALIGNMENTS

RESULT 1

ID	Q9UIR2	PRELIMINARY;	PRT;	265 AA.
AC	Q9UIR2;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	ATP sulfurylase/APS kinase isoform SK2 (EC 2.7.7.4) (Fragment).			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98337975; PubMed=9671738;			
RA	Kurima K., Warman M.L., Krishnan S., Domowicz M., Krueger R.C. Jr.,			
RA	Deyrup A., Schwartz N.B.;			
RT	"A member of a family of sulfate-activating enzymes causes murine			
RT	brachymorphism.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:8681-8685(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20026854; PubMed=10559207;			
RA	Kurima K., Singh B., Schwartz N.B.;			
RT	"Genomic organization of the mouse and human genes encoding the ATP			
RT	Sulfurylase/Adenosine 5'-phosphosulfate kinase isoform SK2.";			
RL	J. Biol. Chem. 274:33306-33312(1999).			
DR	EMBL; AF173364; AAF12780.1;			
DR	EMBL; AF173361; AAF12780.1; JOINED.			
DR	EMBL; AF173362; AAF12780.1; JOINED.			
DR	EMBL; AF173363; AAF12780.1; JOINED.			
DR	InterPro: IPR002650; ATP-sulfurylase.			
DR	Pfam: PF01747; ATP-sulfurylase; 1.			

Q9zwm0 oryza sativ
Q48888 zea mays (m
Q27128 urechis cau
Q96349 brassica na
Q9zn28 brassica ju
Q42520 arabidopsis
Q96530 arabidopsis
Q23324 arabidopsis
Q957d8 arabidopsis
Q9zn29 brassica ju
Q43870 arabidopsis
Q8rwj3 arabidopsis
Q96541 brassica ol
Q9ab10 brassica ol
Q961a8 drosophila
Q9vw48 drosophila
Q95p41 aedes aegypt
Q22501 caenorhabdl
Q9p7g9 schizosacch
Q9se92 zea mays (m
Q9zvh6 rhizobium m
Q985q4 rhizobium l
Q9sqr9 arabidopsis
Q9sqr7 arabidopsis
Q81351 enteromorph
Q9fjx1 arabidopsis
Q9a882 caulobacter
Q8yex3 bruceella me
Q97mt8 clostridium
Q9eyl1 klebsiella
Q8xf34 salmonella
Q9hg18 saccharomyc
Q9c3y6 saccharomyc
Q8zbp3 yersinia pe
Q8rell fusobacteri

DR ProDom: PD002381; ATP-sulphurylase; 1.
KW Kinase: Nucleotidyltransferase; Transferase.
FT NON_TER 1
SQ SEQUENCE 265 AA; 30142 MW; D75A2374E40C334F CRC64;

Query Match 23.18; Score 142; DB 4; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.5e-145;
Matches 142; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 427 RLLRGYKHPVLLHPLGKWTGKDDVPLDWRMKQHAHVLEEGVLDPKSTIVAIFFSPML 486
|||||
Db 70 RLLRGYKHPVLLHPLGKWTGKDDVPLDWRMKQHAHVLEEGVLDPKSTIVAIFFSPML 129
|||||

Qy 487 YAGTEVQWCHRSRMIAAGNFYIVGRDPAGMPHPETKDKLYEPETHGKVLSPMAPGLTSVE 546
|||||
Db 130 YAGTEVQWCHRSRMIAAGNFYIVGRDPAGMPHPETKDKLYEPETHGKVLSPMAPGLTSVE 189
|||||

Qy 547 IIPRVAAYNKKAKAMDFYDPA 568
|||||
Db 190 IIPRVAAYNKKAKAMDFYDPA 211
|||||

RESULT 2
Q9QYS0 PRELIMINARY; PRT; 621 AA.
ID Q9QYS0
AC Q9QYS0
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ATP sulphyrylase/APS kinase isoform SK2 (EC 2.7.7.4).
GN PAPSS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=98337975; PubMed=9671738;
RA Kurima K., Warman M.L., Krishnan S., Domowicz M., Krueger R.C. Jr.,
RA Deyrup A., Schwartz N.B.;
RT "A member of a family of sulfate-activating enzymes causes murine
RT brachymorphism";
RL Proc. Natl. Acad. Sci. U.S.A. 95:8681-8685(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20026854; PubMed=10559207;
RA Kurima K., Singh B., Schwartz N.B.;
RT "Genomic organization of the mouse and human genes encoding the ATP
RT Sulphyrylase/Adenosine 5'-phosphosulfate kinase isoform SK2.";
RL J. Biol. Chem. 274:33306-33312(1999).
DR EMBL; AF172866; AAF12760.1;
DR EMBL; AF172857; AAF12760.1; JOINED.
DR EMBL; AF172858; AAF12760.1; JOINED.
DR EMBL; AF172859; AAF12760.1; JOINED.
DR EMBL; AF172860; AAF12760.1; JOINED.
DR EMBL; AF172861; AAF12760.1; JOINED.
DR EMBL; AF172862; AAF12760.1; JOINED.
DR EMBL; AF172863; AAF12760.1; JOINED.
DR EMBL; AF172864; AAF12760.1; JOINED.
DR EMBL; AF172865; AAF12760.1; JOINED.
DR MGI.1330223; Papss2.
DR InterPro: IPR002891; APS_kinase.
DR InterPro: IPR002650; ATP-sulphyrylase.
DR Pfam; PF01583; APS_kinase; 1.
DR Pfam; PF01747; ATP-sulphyrylase; 1.
DR ProDom: PD002350; APS_kinase; 1.
DR ProDom: PD002381; ATP-sulphyrylase; 1.
DR TIGRFAMS; TIGR00455; apsk; 1.
DR Kinase: Nucleotidyltransferase; Transferase.
SQ SEQUENCE 621 AA; 70408 MW; 019519DC8E239F68 CRC64;

Query Match 13.28; Score 81; DB 11; Length 621;
Best Local Similarity 100.0%; Pred. No. 9.2e-79;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 LFADAGLVCTISFISPPAKORENARKIHESAGLFFFEIFVDAPLNICESRDVKGLYKRR 168
|||||
Db 110 LFADAGLVCTISFISPPAKORENARKIHESAGLFFFEIFVDAPLNICESRDVKGLYKRR 169
|||||

Qy 169 AGEIKGFTGIDSDYEKPEPTE 189
|||||
Db 170 AGEIKGFTGIDSDYEKPEPTE 190
|||||

RESULT 3
Q9JK86 PRELIMINARY; PRT; 620 AA.
ID Q9JK86
AC Q9JK86
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Adenosine 5'-phosphosulfate kinase/ATP sulphyrylase 2.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Fuda H., Strott C.A.;
RT "Guinea pig bifunctional adenosine 5'-phosphosulfate kinase/ATP
RT sulphyrylase 2, gpAPS synthase 2";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF251798; AAF70194.1;
DR InterPro: IPR002891; APS_kinase.
DR InterPro: IPR002650; ATP-sulphyrylase.
DR Pfam; PF01583; APS_kinase; 1.
DR Pfam; PF01747; ATP-sulphyrylase; 1.
DR ProDom: PD002350; APS_kinase; 1.
DR ProDom: PD002381; ATP-sulphyrylase; 1.
DR TIGRFAMS; TIGR00455; apsk; 1.
DR Kinase.
SQ SEQUENCE 620 AA; 70031 MW; 50E187D16335985A CRC64;

Query Match 9.48; Score 58; DB 11; Length 620;
Best Local Similarity 100.0%; Pred. No. 1e-53;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 500 RMIAGANFYIVGRDPAGMPHPETKDKLYEPETHGKVLSPMAPGLTSVEIIPRVAAYNK 557
|||||
Db 505 RMIAGANFYIVGRDPAGMPHPETKDKLYEPETHGKVLSPMAPGLTSVEIIPRVAAYNK 562
|||||

RESULT 4
Q9UE98 PRELIMINARY; PRT; 624 AA.
ID Q9UE98
AC Q9UE98
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ATP sulphyrylase/APS kinase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Deyrup A.T.;
RT "Human ATP sulphyrylase/APS kinase";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF016496; AAD09325.1;
DR InterPro: IPR002891; APS_kinase.
DR InterPro: IPR002650; ATP-sulphyrylase.
DR Pfam; PF01583; APS_kinase; 1.
DR Pfam; PF01747; ATP-sulphyrylase; 1.

DR ProDom; PD002350; APS_kinase; 1.
 DR ProDom; PD002381; ATP-sulphurylase; 1.
 DR TIGRFAMS; TIGR00455; apsk; 1.
 KW Kinase.
 SQ SEQUENCE 624 AA; 70881 MW; 0008DBC57B6F35BC CRC64;
 Query Match 6.2%; Score 38; DB 4; Length 624;
 Best Local Similarity 100.0%; Pred. No. 6.2e-32;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 57
 Db 30 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 67
 RESULT 5
 Q9P1P9
 ID Q9P1P9 PRELIMINARY; PRT; 624 AA.
 AC Q9P1P9
 DT 01-OCT-2000 (TremBLrel. 15, Created)
 DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE 3'-phosphoadenosine 5'-phosphosulfate synthetase.
 GN PAPSS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=20145452; PubMed=10679223;
 RA Xu Z.-H., Otterness D.M., Freimuth R.R., Carlini E.J., Wood T.C.,
 RA Mitchell S., Moon E., Kim U.-J., Xu J.-P., Siciliano M.J.,
 RA Weinshilboum R.M.;
 RT PAPSS2: gene cloning, characterization and chromosomal localization.;
 RL Biochem. Biophys. Res. Commun. 268:437-444(2000).
 DR EMBL; AF105227; AAF40236.1; JOINED.
 DR InterPro; IPR002891; APS_kinase.
 DR Pfam; PF01583; APS_kinase; 1.
 DR ProDom; PD002350; APS_kinase; 1.
 DR ProDom; PD002381; ATP-sulphurylase; 1.
 DR TIGRFAMS; TIGR00455; apsk; 1.
 SQ SEQUENCE 624 AA; 70859 MW; 60DC9B943E7B75ED CRC64;
 Query Match 6.2%; Score 38; DB 4; Length 624;
 Best Local Similarity 100.0%; Pred. No. 6.2e-32;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 57
 Db 30 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 67
 RESULT 6
 Q96TF4
 ID Q96TF4 PRELIMINARY; PRT; 624 AA.
 AC Q96TF4; Q96FBI;
 DT 01-DEC-2001 (TremBLrel. 19, Created)
 DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE 3'-phosphoadenosine 5'-phosphosulfate synthetase.
 GN PAPSS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20145452; PubMed=10679223;

RA Xu Z.-H., Otterness D.M., Freimuth R.R., Carlini E.J., Wood T.C.,
 RA Mitchell S., Moon E., Kim U.-J., Xu J.-P., Siciliano M.J.,
 RA Weinshilboum R.M.;
 RT "Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSS1) and
 RT PAPSS2: gene cloning, characterization and chromosomal localization.";
 RL Biochem. Biophys. Res. Commun. 268:437-444(2000).
 RN [2]
 RN SEQUENCE OF 22-624 FROM N.A.
 RC TISSUE=EYE;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; AF097721; AAF40235.1; JOINED.
 DR EMBL; AF097710; AAF40235.1; JOINED.
 DR EMBL; AF097711; AAF40235.1; JOINED.
 DR EMBL; AF097712; AAF40235.1; JOINED.
 DR EMBL; AF097713; AAF40235.1; JOINED.
 DR EMBL; AF097714; AAF40235.1; JOINED.
 DR EMBL; AF097715; AAF40235.1; JOINED.
 DR EMBL; AF097716; AAF40235.1; JOINED.
 DR EMBL; AF097717; AAF40235.1; JOINED.
 DR EMBL; AF097718; AAF40235.1; JOINED.
 DR EMBL; AF097719; AAF40235.1; JOINED.
 DR EMBL; AF097720; AAF40235.1; JOINED.
 DR EMBL; BC011392; AAH11392.1; JOINED.
 DR InterPro; IPR002891; APS_kinase.
 DR InterPro; IPR002650; ATP-sulphurylase.
 DR Pfam; PF01583; APS_kinase; 1.
 DR Pfam; PF01747; ATP-sulphurylase; 1.
 DR ProDom; PD002350; APS_kinase; 1.
 DR ProDom; PD002381; ATP-sulphurylase; 1.
 DR TIGRFAMS; TIGR00455; apsk; 1.
 SQ SEQUENCE 624 AA; 70833 MW; A3DC9B943E68CDD6 CRC64;
 Query Match 6.2%; Score 38; DB 4; Length 624;
 Best Local Similarity 100.0%; Pred. No. 6.2e-32;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 57
 Db 30 YQAHVSRNKRQGVVTRGGFGCTVWLTGLSGAGKTT 67
 RESULT 7
 Q90XY2
 ID Q90XY2 PRELIMINARY; PRT; 613 AA.
 AC Q90XY2
 DT 01-DEC-2001 (TremBLrel. 19, Created)
 DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE 3'-phosphoadenosine 5'-phosphosulfate synthase 2.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Takifugu.
 OX NCBI_TaxID=31033;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=21455682; PubMed=11571655;
 RA Yu W.P., Pallen C.J., Tay A., Jirik F.R., Brenner S., Tan Y.H.,
 RA Venkatesh B.;
 RT "Conserved synteny between the Fugu and human PTEN locus and the
 RT evolutionary conservation of vertebrate PTEN function.";
 RL Oncogene 20:5554-5561(2001).
 DR EMBL; AF325922; AAL08416.1; JOINED.
 DR InterPro; IPR002891; APS_kinase.
 DR InterPro; IPR002650; ATP-sulphurylase.
 DR Pfam; PF01583; APS_kinase; 1.
 DR Pfam; PF01747; ATP-sulphurylase; 1.
 DR ProDom; PD002350; APS_kinase; 1.
 DR ProDom; PD002381; ATP-sulphurylase; 1.
 DR TIGRFAMS; TIGR00455; apsk; 1.
 SQ SEQUENCE 613 AA; 68948 MW; 28E50148377C4169 CRC64;

Query Match 6.0%; Score 37; DB 13; Length 613;
 Best Local Similarity 100.0%; Pred. No. 7.4e-31;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 PVLLHPLGGTWDVPLDWRMKOHAHVLEEGVLD 473
 Db 436 PVLLHPLGGTWDVPLDWRMKOHAHVLEEGVLD 472

RESULT 8

Q9NDP8 ID Q9NDP8 PRELIMINARY; PRT; 618 AA.
 AC Q9NDP8; 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ATP sulfurylase/APS kinase.
 GN CI-ASAK.
 OS Clona intestinalis.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 OC Phlebobranchia; Clonidae; Clona.
 OX NCBI_TaxID=7719;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hotta K., Takahashi H., Asakura T., Saitoh B., Takatori N., Satou Y.,
 RA Satou N.;
 RT "Characterization of Brachyury downstream notochord genes in the Clona
 RL intestinalis embryo."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB036852; BAB00629.1; -
 DR InterPro: IPR002891; APS_kinase.
 DR InterPro: IPR002650; ATP-sulfurylase.
 DR Pfam: PF01583; APS_kinase; 1.
 DR ProDom: PD002350; APS_kinase; 1.
 DR ProDom: PD002381; ATP-sulfurylase; 1.
 DR TIGRFAMs: TIGR00455; apsk; 1.
 KW Kinase.
 SQ SEQUENCE 618 AA; 69559 MW; 75223CCB02BC397 CRC64;

Query Match 4.2%; Score 26; DB 5; Length 618;
 Best Local Similarity 100.0%; Pred. No. 7.1e-19;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 400 MNADAVFAFQLRNPVHNGHALLMQDT 425
 Db 403 MNADAVFAFQLRNPVHNGHALLMQDT 428

RESULT 9

Q43170 ID Q43170 PRELIMINARY; PRT; 424 AA.
 AC Q43170;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Sulfate adenylyltransferase (EC 2.7.7.4).
 GN STMET3-1.
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DESIREE; TISSUE=LEAF;
 RX MEDLINE=95004649; PubMed=7920699;
 RA Klonous D., Hoergen R., Willmitzer L., Riesmeier J.W.;
 RT "Isolation and characterization of two cDNAs clones encoding ATP-
 RT sulfurylase from Potato by complementation of a yeast mutant."
 RL Plant J. 6:105-112(1994).
 DR EMBL: X75041; CAA52953.1; -

DR InterPro: IPR002650; ATP-sulfurylase.
 DR Pfam: PF01747; ATP-sulfurylase; 1.
 DR ProDom: PD002381; ATP-sulfurylase; 1.
 DR TIGRFAMs: TIGR00339; sopt; 1.
 KW Nucleotidyltransferase; Transferase.
 SQ SEQUENCE 424 AA; 47519 MW; E82A27DC11350ABC CRC64;

Query Match 3.6%; Score 22; DB 10; Length 424;
 Best Local Similarity 100.0%; Pred. No. 1.1e-14;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAFQLRNPVHNGHALLM 422
 Db 200 NADAVFAFQLRNPVHNGHALLM 221

RESULT 10

O8W1X3 ID O8W1X3 PRELIMINARY; PRT; 458 AA.
 AC O8W1X3;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ATP-sulfurylase.
 OS Allium cepa (Onion).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
 OC Allium.
 OX NCBI_TaxID=4679;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. PUKEKOHE LONG KEEPER; TISSUE=ROOT;
 RA Pither-Joyce M.D., McCallum J.A.;
 RT "Isolation of an ATP-sulfurylase cDNA clone from Allium cepa."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF403295; AAL61615.1; -
 DR InterPro: IPR002650; ATP-sulfurylase.
 DR Pfam: PF01747; ATP-sulfurylase; 1.
 DR ProDom: PD002381; ATP-sulfurylase; 1.
 DR TIGRFAMs: TIGR00339; sopt; 1.
 SQ SEQUENCE 458 AA; 51360 MW; F7A43A53F3A7F04C CRC64;

Query Match 3.6%; Score 22; DB 10; Length 458;
 Best Local Similarity 100.0%; Pred. No. 1.2e-14;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 401 NADAVFAFQLRNPVHNGHALLM 422
 Db 233 NADAVFAFQLRNPVHNGHALLM 254

RESULT 11

Q9SDP4 ID Q9SDP4 PRELIMINARY; PRT; 461 AA.
 AC Q9SDP4;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ATP-sulfurylase.
 OS Allium cepa (Onion).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
 OC Allium.
 OX NCBI_TaxID=4679;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. PUKEKOHE LONG KEEPER; TISSUE=ROOT;
 RA Pither-Joyce M.D., McCallum J.A.;
 RT "Isolation of an ATP-sulfurylase cDNA clone from Allium cepa."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF212154; AAF18998.1; -
 DR InterPro: IPR002650; ATP-sulfurylase.
 DR Pfam: PF01747; ATP-sulfurylase; 1.

DR ProDom; PD002381; ATP-sulfurylase; 1.
DR TIGRFAMS; TIGR00339; sopt; 1.
SQ SEQUENCE 461 AA; 51807 MW; 04083BE879CBDE71 CRC64;

Query Match 3.6%; Score 22; DB 10; Length 461;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVFAFQLRNPVNHGHALLM 422
|||||

Db 236 NADAVFAFQLRNPVNHGHALLM 257

RESULT 12

Q8SAG1 PRELIMINARY; PRT; 461 AA.

AC Q8SAG1;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ATP sulfurylase.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SEEDLING;
RA Phartival P., Krishnan H.B.;
RT "Nucleotide sequence of a soybean cDNA encoding adenosine triphosphate
sulfurylase";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF452454; AAL74418.1; -;
SQ SEQUENCE 461 AA; 51700 MW; 3504272C93923100 CRC64;

Query Match 3.6%; Score 22; DB 10; Length 461;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVFAFQLRNPVNHGHALLM 422
|||||

Db 238 NADAVFAFQLRNPVNHGHALLM 259

RESULT 13

Q42519 PRELIMINARY; PRT; 463 AA.

AC Q42519;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ATP sulfurylase precursor (EC 2.7.7.4).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=9433677; PubMed=8058839;
RA Leustek T., Murillo M., Cervantes M.;
RT "Cloning of a cDNA encoding ATP sulfurylase from Arabidopsis thaliana
by functional expression in Saccharomyces cerevisiae";
RL Plant Physiol. 105:897-902(1994).
DR EMBL; U05218; AAA21570.1; -;
DR InterPro; IPR002650; ATP-sulfurylase.
DR Pfam; PF01747; ATP-sulfurylase; 1.
DR ProDom; PD002381; ATP-sulfurylase; 1.
DR TIGRFAMS; TIGR00339; sopt; 1.
KW Chloroplast; Nucleotidyltransferase; Transferase; Transit peptide.
FT TRANSIT 1 48 CHLOROPLAST.
FT CHAIN 49 463 ATP SULFURYLASE.

SQ SEQUENCE 463 AA; 51372 MW; 9A45FAA5133A17FF CRC64;

Query Match 3.6%; Score 22; DB 10; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVFAFQLRNPVNHGHALLM 422
|||||

Db 239 NADAVFAFQLRNPVNHGHALLM 260

RESULT 14

Q9SE02 PRELIMINARY; PRT; 463 AA.

AC Q9SE02;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE ATP sulfurylase (EC 2.7.7.4).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Hatzfeld Y., Lee S., Lee M., Leustek T., Saito K.;
RT "Functional characterization of a gene encoding a fourth ATP
sulfurylase isoform from Arabidopsis thaliana";
RL Gene 248:51-58(2000).
DR EMBL; AF198964; AAF19185.1; -;
DR InterPro; IPR002650; ATP-sulfurylase.
DR Pfam; PF01747; ATP-sulfurylase; 1.
DR ProDom; PD002381; ATP-sulfurylase; 1.
DR TIGRFAMS; TIGR00339; sopt; 1.
DR Nucleotidyltransferase; Transferase.
KW SEQUENCE 463 AA; 51444 MW; 18EC241A3B46CB3D CRC64;

Query Match 3.6%; Score 22; DB 10; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 401 NADAVFAFQLRNPVNHGHALLM 422
|||||

Db 239 NADAVFAFQLRNPVNHGHALLM 260

RESULT 15

Q43183 PRELIMINARY; PRT; 463 AA.

AC Q43183;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Sulfate adenylyltransferase (EC 2.7.7.4).
GN MET3-2.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DESIREE; TISSUE=LEAF;
RA MEDLINE=95004649; PubMed=7920699;
RA Klonus D., Hoefgen R., Willmitzer L., Riesmeier J.W.;
RT "Isolation and characterization of two cDNAs encoding ATP-
sulfurylase from potato by complementation of a yeast mutant";
RL Plant J. 6:105-112(1994).
DR EMBL; X79053; CAA55655.1; -;
DR InterPro; IPR002650; ATP-sulfurylase.

Query Match 3.6%; Score 22; DB 10; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.2e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: May 30, 2003, 12:32:18
Job time : 89 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:08:55 ; Search time 73 Seconds
(without alignments)
1122.591 Million cell updates/sec

Title: US-09-898-165B-7

Perfect score: 3271

Sequence: 1 MSGIKKQKTKNQKSTNVY.....MAPKAWKLVTDYRSEMDKN 615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	3216	98.3	619	21 AAY79214	Human transferase
2	3162.5	96.7	628	22 AAU32438	Novel human secret
3	2602	79.5	624	19 AAW70494	Human disease rela
4	2602	79.5	624	20 AAY22349	Human APS kinase/A
5	2602	79.5	625	20 AAW67882	Human secreted pro
6	2031.5	62.1	635	22 ABB71650	Drosophila melanog
7	1294	39.6	460	21 AAG44040	Arabidopsis thalia
8	1294	39.6	463	21 AAG44039	Arabidopsis thalia
9	1282	39.2	462	21 AAG30254	Arabidopsis thalia
10	1282	39.2	465	21 AAG30253	Arabidopsis thalia

11	1282	39.2	510	21 AAG30252	Arabidopsis thalia
12	1272.5	38.9	473	21 AAG09595	Arabidopsis thalia
13	1272.5	38.9	476	21 AAG09594	Arabidopsis thalia
14	1270	38.8	425	21 AAG09596	Arabidopsis thalia
15	1259.5	38.5	408	21 AAG43667	Arabidopsis thalia
16	1259.5	38.5	469	21 AAG43666	Arabidopsis thalia
17	1252.5	38.3	396	21 AAG43668	Arabidopsis thalia
18	1228	37.5	373	21 AAG44041	Arabidopsis thalia
19	1225.5	37.5	446	21 AAG21121	Arabidopsis thalia
20	1223.5	37.4	385	21 AAG21122	Arabidopsis thalia
21	1216.5	37.2	373	21 AAG21123	Arabidopsis thalia
22	616	18.8	174	21 AAG20074	Arabidopsis thalia
23	581	17.8	166	21 AAG20075	Arabidopsis thalia
24	508.5	15.5	311	21 AAY44791	Soybean Adenylisulph
25	499.5	15.3	246	21 AAY44792	Wheat Adenylisulph
26	497.5	15.2	293	21 AAG29425	Arabidopsis thalia
27	497.5	15.2	293	21 AAY77957	A. thaliana enviro
28	482	14.7	224	21 AAY44788	Corn Adenylisulph
29	482	14.7	343	21 AAY44789	Corn Adenylisulph
30	481	14.7	204	23 ABP39672	Staphylococcus epi
31	476.5	14.6	201	22 AAG98409	Escherichia coli p
32	476	14.6	236	21 AAG14478	Arabidopsis thalia
33	476	14.6	251	21 AAG11017	Arabidopsis thalia
34	476	14.6	251	21 AAG38348	Arabidopsis thalia
35	476	14.6	259	21 AAG14477	Arabidopsis thalia
36	476	14.6	263	21 AAG11016	Arabidopsis thalia
37	476	14.6	263	21 AAG38347	Arabidopsis thalia
38	476	14.6	272	21 AAG14476	Arabidopsis thalia
39	476	14.6	276	21 AAG11015	Arabidopsis thalia
40	476	14.6	276	21 AAG38346	Arabidopsis thalia
41	472.5	14.4	208	21 AAG35788	Arabidopsis thalia
42	472	14.4	140	21 AAG20076	Arabidopsis thalia
43	458.5	14.0	252	21 AAG14580	Arabidopsis thalia
44	458.5	14.0	252	21 AAG47458	Arabidopsis thalia
45	458.5	14.0	305	21 AAG14579	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AAV79214

ID AAY79214 standard; Protein; 619 AA.

XX AAY79214;

XX AC AAY79214;

XX DT 19-JUN-2000 (first entry)

XX DE Human transferase TRNSFS-6.

XX KW Transferase; TRNSFS-6; human; antitumour; cancer;

XX KW gastrointestinal disorder; developmental disorder;

XX KW genetic disorder; neurological disorder; reproductive disorder;

XX KW smooth muscle disorder; immunological disorder; inflammation;

XX KW diagnosis; therapy; ATP sulfurylase/APS kinase 2.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 92

FT Modified-site 176

FT Modified-site 180

FT Modified-site 198

FT Modified-site 250

FT Modified-site 285

FT Modified-site 308

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site /note= "potential O-phosphorylation"

FT Modified-site /note= "potential O-phosphorylation"

Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

QY 5 KKOKTEN-----QQKSTNVYQAHHVSRNRKRGQVGTGRGCTVWLTGLSGAGKTTI 58
 DB 9 KKVLSNNAQNWGMQRATNTYQAHHVSRNRKRGQVGTGRGCTVWLTGLSGAGKTTV 68

QY 59 SFALIEEYLVSHAIPCYSLDGDNVHGLNRNLGSPGDRREENIRIARIVAKLFDAGLVCI 118
 DB 69 SMALEEYLVCHGIPCYTLDDGNIHQGLNKLGFSPEDREENVRRIAEVAKLFDAGLVCI 128

QY 119 TSFISPAKRENAKITHESAGLPFFEIFVDAPLNICESDVGLYKRRARAGEIKGFTGI 178
 DB 129 TSFISPTQDRNARQIHGASLPFFEFVFDAPLHCEQDRVGLYKRRARAGEIKGFTGI 188

QY 179 DSYEKPTPERVLKTNLSTVSDCVHQVVELLOEQNIIVPTIHKDIHELFVPEKNKLDHVR 238
 DB 189 DSEYKPEAPELVKTDSCDVNDVQVQVVELLOEQNIIVPTIHKDIHELFVPEKNKLDHVR 248

QY 239 AEATLPSLITKLDLQWQVLSGWTPLKGFMRKEYLOVMHFDTLDDGVINLSVPI 298
 DB 249 TDAETLPALKINKVDQWQVLAEGWATPLNGFMREYELQCLHFDCLDDGVINLSVPI 308

QY 299 VLPVSAEDKRLGECSCFVLAHGRVAILRDADAEFEYHRRKEERCARQWGTTCCKHPIKM 358
 DB 309 VLTATHEDKERLOGCTAFALMYEGRRVAILRNPEFEHRRKEERCARQWGTTCCKHPIKM 368

QY 359 VMESGDLVGGDLQVLEKIRWNDGLDQYRLTPELKKOKKEMNADAVFAFQLRNPVHNGH 418
 DB 369 VMEQGDWLIGDGLQVLDVRYWVNDGLDQYRLTPELKKOKKEMNADAVFAFQLRNPVHNGH 428

QY 419 ALLMODTCRLLERGVKHPVLLHPLGGWTKDDVPLDWRMKOHAALVEEGVLDPKSTIV 478
 DB 429 ALLMODTHKOLLERGVRRPVLLHPLGGWTKDDVPLDWRMKOHAALVEEGVLDPKSTIV 488

QY 479 AIFPSPMLYAGPTEVQWCHCRSMIAGANFYIVGRDPAGMPHPTKTDLYEPTTHGGKVLMS 538
 DB 489 AIFPSPMAYAGPTEVQWCHCRARMVAGANFYIVGRDPAGMPHPTKTDLYEPTTHGGKVLMS 548

QY 539 APGLTSEIIPFRVAAYNKAAMDFYDPAHNEFDIFISGTRMKRLAREGENPPDGMFMAP 598
 DB 549 APGLITLIEVPRVAAAYNKKRMDDYDSEHHEDEFFISGTRMKRLAREQKPPGEGFMAP 608

QY 599 KAWKVLTDYKRS 610
 DB 609 KAWTVLLEYKRS 620

RESULT 4
 AAY22349
 ID AAY22349 standard; Protein; 624 AA.

AC AAY22349;

DT 24-SEP-1999 (first entry)

DE Human APS kinase/ATP sulphurylase protein.

XX APS kinase/ATP sulphurylase; human; PAPS production;
 3'-phosphoadenosine 5'-phosphosulphate.

OS Homo sapiens.

PN JF11187883-A.

PD 13-JUL-1999.

PF 26-DEC-1997; 97JP-0360387.

XX 26-DEC-1997; 97JP-0360387.

XX (NIRA) UNITIKA LTD.

PA (HUMA-) ZH HUMAN SCI SHINKO ZAIDAN.

XX

DR WPI; 1999-451549/38.
 DR N-PSDB; AAX84897.
 XX New human-derived APS kinase/ATP sulphyrylase gene - useful for
 PT large scale preparation of 3'-phosphoadenosine 5'-phosphosulfate
 PT (PAPS)
 XX Claim 1; Page 6-7; 9pp; Japanese.
 PS This sequence is the human-derived APS kinase/ATP sulphyrylase of
 CC the invention. The enzyme may be used to prepare 3'-phosphoadenosine
 CC 5'-phosphosulphate (PAPS) on a large scale.
 XX

SQ Sequence 624 AA;
 Query Match 79.5%; Score 2602; DB 20; Length 624;
 Best Local Similarity 77.1%; Pred. No. 4.2e-254;
 Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

QY 5 KKOKTEN-----QQKSTNVYQAHHVSRNRKRGQVGTGRGCTVWLTGLSGAGKTTI 58
 DB 9 KKVLSNNAQNWGMQRATNTYQAHHVSRNRKRGQVGTGRGCTVWLTGLSGAGKTTV 68

QY 59 SFALIEEYLVSHAIPCYSLDGDNVHGLNRNLGSPGDRREENIRIARIVAKLFDAGLVCI 118
 DB 69 SMALEEYLVCHGIPCYTLDDGNIHQGLNKLGFSPEDREENVRRIAEVAKLFDAGLVCI 128

QY 119 TSFISPAKRENAKITHESAGLPFFEIFVDAPLNICESDVGLYKRRARAGEIKGFTGI 178
 DB 129 TSFISPTQDRNARQIHGASLPFFEFVFDAPLHCEQDRVGLYKRRARAGEIKGFTGI 188

QY 179 DSYEKPTPERVLKTNLSTVSDCVHQVVELLOEQNIIVPTIHKDIHELFVPEKNKLDHVR 238
 DB 189 DSEYKPEAPELVKTDSCDVNDVQVQVVELLOEQNIIVPTIHKDIHELFVPEKNKLDHVR 248

QY 239 AEATLPSLITKLDLQWQVLSGWTPLKGFMRKEYLOVMHFDTLDDGVINLSVPI 298
 DB 249 TDAETLPALKINKVDQWQVLAEGWATPLNGFMREYELQCLHFDCLDDGVINLSVPI 308

QY 299 VLPVSAEDKRLGECSCFVLAHGRVAILRDADAEFEYHRRKEERCARQWGTTCCKHPIKM 358
 DB 309 VLTATHEDKERLOGCTAFALMYEGRRVAILRNPEFEHRRKEERCARQWGTTCCKHPIKM 368

QY 359 VMESGDLVGGDLQVLEKIRWNDGLDQYRLTPELKKOKKEMNADAVFAFQLRNPVHNGH 418
 DB 369 VMEQGDWLIGDGLQVLDVRYWVNDGLDQYRLTPELKKOKKEMNADAVFAFQLRNPVHNGH 428

QY 419 ALLMODTCRLLERGVKHPVLLHPLGGWTKDDVPLDWRMKOHAALVEEGVLDPKSTIV 478
 DB 429 ALLMODTHKOLLERGVRRPVLLHPLGGWTKDDVPLDWRMKOHAALVEEGVLDPKSTIV 488

QY 479 AIFPSPMLYAGPTEVQWCHCRSMIAGANFYIVGRDPAGMPHPTKTDLYEPTTHGGKVLMS 538
 DB 489 AIFPSPMAYAGPTEVQWCHCRARMVAGANFYIVGRDPAGMPHPTKTDLYEPTTHGGKVLMS 548

QY 539 APGLTSEIIPFRVAAYNKAAMDFYDPAHNEFDIFISGTRMKRLAREGENPPDGMFMAP 598
 DB 549 APGLITLIEVPRVAAAYNKKRMDDYDSEHHEDEFFISGTRMKRLAREQKPPGEGFMAP 608

QY 599 KAWKVLTDYKRS 610
 DB 609 KAWTVLLEYKRS 620

RESULT 5
 AAW67882
 ID AAW67882 standard; Protein; 625 AA.

XX AAW67882;

XX 25-MAR-1999 (first entry)

DE Human secreted protein encoded by gene 76 clone HOSFD58.

XX Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX Homo sapiens.

OS
XX
FH Key Location/Qualifiers
FT Misc-difference 625
FT /label= unknown
XX

PN WO9842738-A1.

PD 01-OCT-1998.

XX 19-MAR-1998; 98WO-US05311.

PF 30-MAY-1997; 97US-0050937.

PR 21-MAR-1997; 97US-0041276.

PR 21-MAR-1997; 97US-0041277.

PR 21-MAR-1997; 97US-0041281.

PR 21-MAR-1997; 97US-0042344.

PR 30-MAY-1997; 97US-0048069.

PR 30-MAY-1997; 97US-0048094.

PR 30-MAY-1997; 97US-0048095.

PR 30-MAY-1997; 97US-0048096.

PR 30-MAY-1997; 97US-0048099.

PR 30-MAY-1997; 97US-0048131.

PR 30-MAY-1997; 97US-0048135.

PR 30-MAY-1997; 97US-0048154.

PR 30-MAY-1997; 97US-0048160.

PR 30-MAY-1997; 97US-0048186.

PR 30-MAY-1997; 97US-0048187.

PR 30-MAY-1997; 97US-0048188.

PR 30-MAY-1997; 97US-0048350.

PR 30-MAY-1997; 97US-0048351.

PR 30-MAY-1997; 97US-0048352.

PR 30-MAY-1997; 97US-0048355.

PR 05-AUG-1997; 97US-0054804.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Brewer LA, Duan R, Ebner R, Ferrie AM, Florence KA;

PI Greene JM, Hu JS, Lafleur DW, Moore PA, Ni J, Olsen HS;

PI Rosen CA, Ruben SM, Shi Y, Young P;

XX WPI: 1999-070066/06.

DR N-PSDB; AAX00686.

CC (see AAX00611 for described uses).

XX Sequence 625 AA;

Query Match 79.5%; Score 2602; DB 20; Length 625;

Best Local Similarity 77.1%; Pred. No. 4.2e-254;

Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

QY 5 KKQKTE-----QOKSTNVVYQAHHSVRNKRKGVVGRGFRGCTVWLTGLSGAGKTTI 58

DB 9 KKVLSNNAQNWGMQRATNVYQAHHSVRNKRKGVVGRGFRGCTVWLTGLSGAGKTTV 68

QY 59 SFALBEYLVSRAIPCYSLDGDNRVHGLNRNLGFSFGDREENTRRRTAEVAKLADAGLVCI 118

DB 69 SMALEELVCHGICPYCTLDGNIROGLNKNLGFSPEDREENVYRRTAEVAKLADAGLVCI 128

QY 119 TSFISPFADKRENARKIHESAGLPFFEIFVDAPLNICESRDYKGLYKRARAGEIKGFTGI 178

DB 129 TSFISPYTQDRNNARQIHGASLPFFFEVVDAPLVHCEORDYKGLYKRARAGEIKGFTGI 188

QY 179 DSDYEKPEPVRVLTNLSVSDCVHVQVVELLEQONIVPYTTIKDIHELFPVENKLDHVR 238

DB 189 DSEYEKPEAPELVLTDSQVNDVQVVELLQERDIPVDASYEVKELYVPENKLHLAK 248

QY 239 AEAETLPSLITKLDLQWVQLSEGWATPLKGFMEKEKYLQVWHEDTLLDDGVNMSIPI 298

DB 249 TDAETLPALKINKVDMQWQVLAEGWATPLNGFREREYLOCLHFDCLLDGGVINLSVPI 308

QY 299 VLPVSAEDKTRLEGCKFVLAHGRVRVAILRDAEFYEHRKEKRSRVWGTCTTKHPHIKM 358

DB 309 VLTATHEDKERLDGCTAFALMYEGRRVAILRNPFEFEHRKEERCAQWGTCKNHPYIKM 368

QY 359 VNESGDMVLGGDLQVLEKIRWNGDLQVRLTPLELKQCKENNAVAFQLRNPNVHNGH 418

DB 369 VMEQDMLTGGDLQVLDVRYWNGDLQVRLTPELTKQKFKDMNADAVFAFQLRNPNVHNGH 428

QY 419 ALLMODTCRRLERCKYKHPVLLHPLGGWTKDDVPLDWRMKQHAALVEEGVLDPKSTIV 478

DB 429 ALLMODTHKQLLERGRYRRPVLHPLGGWTKDDVPLDWRMKQHAALVEEGVLPNPTVV 488

QY 479 AIFPSPMLYAGPTEVQWCHRCRSMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLSM 538

DB 489 AIFPSPMLYAGPTEVQWCHRCRSMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLSM 548

QY 539 APLTSVEIIPRVAAYNKKAKKAMFYDPAHNEFDFTSGTMRKLAREGENPPDGFMAP 598

DB 549 APLTLTLEIVPRVAAYNKKKRMNDYDSEHEDFEFTSGTMRKLAREGQKPPGFMAP 608

QY 599 KAWKVLTDYRS 610

DB 609 KAWTVLTYEYKS 620

RESULT 6

ABB71650

ID ABB71650 standard; Protein; 635 AA.

XX ABB71650;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 41742.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

OS Drosophila melanogaster.

XX

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX New isolated human genes and the secreted polypeptides they encode -
KW useful for diagnosis and treatment of e.g. cancers, neurological
KW disorders, immune diseases, inflammation or blood disorders
XX Claim 11; Page 313-315; 385pp; English.
XX This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. AAX00602) for increasing the stability of the fused
CC protein as compared to the human protein only.
CC The invention relates to 87 novel genes and their fragments (nucleic
CC acid sequences: AAX00611-X00724; amino acid sequences AAW67807-W68004)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 87
CC polynucleotides, based on which tissues they are most highly expressed in

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139730.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147036.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 39.6%; Score 1294; DB 21; Length 460;

Best Local Similarity 62.2%; Pred. No. 8.8e-122;

Matches 242; Conservative 61; Mismatches 82; Indels 4; Gaps 4;

QY 226 ELFPENKLDHVRRAEAEPLSLITKLDLQWQVLSEGWATPLKGMREKEYLQVMHFD 285

DB 57 ELIVEEPKRREKKHEADLPVELTAIDLQNMHVLSEGWATPLKGMRESEFLQTLHNS 116

QY 286 L-LDDG-VINMSIPVLVPSAEDTRLECCSKFVLHG-GRRVAILRDAEYEHKERC 342

```
Db      117  LRLDDGVSVMNSPVLVAIDDEQKARIGESTRVALENSDGNPVSILSDIEIYKHPKEERI 176
Qy      343  SRVWGTCTTKPHI-KVMESGOWLVGGDLQVLEKIRWNGDLDQYRLTPLELKQCKEMN 401
Db      177  ARTWGTAPGLPVDEAITWAGNWLIGGDLVLEPVKYNIDGLDRFLSPALRKLELEKN 236
Qy      402  ADAVEAFQLRPNVHNGHALLMDTCRRLLRGYKHPVLLHPLGGWTKDDVPLDWRMKQ 461
Db      237  ADAVEAFQLRPNVHNGHALLMDTCRRLLRGYKHPVLLHPLGGWTKDDVPLDWRMKQ 296
Qy      462  HAAVLCEGLDPKSTIVAIPPSPMLYAGPTEVQWCHRSRMAGANFYIVGRDPAGMPHPE 521
Db      297  HEKVLDEGVLDPEPTTVYSIPSPMHYAGPTEVQWCHRSRMAGANFYIVGRDPAGMPH 356
Qy      522  TKKDLVEPTHGGKVLNMAPGLTSVEIIPRVAAYNKAAMDFYDPAHNEPFDISGTRM 581
Db      357  EKRDLDVADHGKVLNMAPGLERINILPFRVAAYDKTQGMKAFDFSRPQDFLFISGTRM 416
Qy      582  RKLAREGENPPDGFMAPKAMKVLTDYYS 610
Db      417  RTLAKNNEPDPGFMCPGGKVLVDYVES 445

RESULT 8
AAG44039
ID      AAG44039 standard; Protein: 463 AA.
AC
AC      AAG44039;
DT
DT      18-OCT-2000 (first entry)
XX
XX      Arabidopsis thaliana protein fragment SEQ ID NO: 55115.
KW      Protein identification; signal transduction pathway; metabolic pathway;
KW      hybridisation assay; genetic mapping; gene expression control; promoter;
KW      termination sequence.
XX
XX      Arabidopsis thaliana.
OS
OS      EP1033405-A2.
PN
PN      06-SEP-2000.
PD
PD      25-FEB-2000; 2000EP-0301439.
XX
XX      25-FEB-1999; 99US-0121825.
XX      05-MAR-1999; 99US-0123180.
XX      09-MAR-1999; 99US-0123548.
XX      23-MAR-1999; 99US-0125788.
XX      25-MAR-1999; 99US-0126264.
XX      29-MAR-1999; 99US-0126785.
XX      01-APR-1999; 99US-0127462.
XX      06-APR-1999; 99US-0128234.
XX      08-APR-1999; 99US-0128714.
XX      16-APR-1999; 99US-0129845.
XX      19-APR-1999; 99US-0130077.
XX      21-APR-1999; 99US-0130449.
XX      23-APR-1999; 99US-0130510.
XX      28-APR-1999; 99US-0130891.
XX      30-APR-1999; 99US-0131449.
XX      30-APR-1999; 99US-0132048.
XX      04-MAY-1999; 99US-0132484.
XX      05-MAY-1999; 99US-0132485.
XX      06-MAY-1999; 99US-0132486.
XX      07-MAY-1999; 99US-0132487.
XX      11-MAY-1999; 99US-0132863.
XX      14-MAY-1999; 99US-0134256.
XX      14-MAY-1999; 99US-0134218.
XX      14-MAY-1999; 99US-0134219.
XX      14-MAY-1999; 99US-0134221.
XX      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144332.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144814.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      23-JUL-1999; 99US-0145145.
PR      23-JUL-1999; 99US-0145218.
PR      23-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
```

PR	26-OCT-1999;	99US-01613559;
PR	26-OCT-1999;	99US-01613601;
PR	26-OCT-1999;	99US-01613611;
PR	28-OCT-1999;	99US-01619201;
PR	28-OCT-1999;	99US-01619202;
PR	28-OCT-1999;	99US-01619931;
PR	28-OCT-1999;	99US-01621421;
PR	29-OCT-1999;	99US-01621422;
Query Match		
Best Local Similarity		39.6%;
Matches 242;		Conservative 62.2%;
QY	226	ELFVPENKLDHVRRAEATL
DB	60	ELIVEPKRREKKHEAADL
QY	286	L-LDDG-VTNMSIPVLVDV
DB	120	LRLLDGSVYNMSVPVLAI
QY	343	SRVYGTCTKKHPHI-KMYVM
DB	180	ARTWGTATGLPYVDEAIT
QY	402	ADAVFAFQLRNVPVNHGHAL
DB	240	ADAVFAFQLRNVPVNHGHAL
QY	462	HAALVEEGVLDPKSTIVAI
DB	300	HEKVLVDGVLDPETTVISL
QY	522	TKKLDYEPHTHGKVLMSAP
DB	360	EKRLDYADAGHKKVLMSAP
QY	582	RKLAREGNPPDGFMAPKAK
DB	420	RTLAKNNENPPDGFMCPCGG
RESULT 9		
AAG30254		
ID	AAG30254	standard; Protein;
XX	AC	
XX	CC	
XX	AC	AAG30254;
DT	DT	
XX	XX	17-OCT-2000 (first entry)
XX	XX	Arabidopsis thaliana protein
XX	XX	Protein identification; signifi-
XX	XX	cation; genetic
XX	XX	terminisation assay; genetic
XX	XX	terminisation sequence.
XX	QS	Arabidopsis thaliana.
XX	XX	
PN	EP1033405-A2.	
XX	XX	
PD	XX	06-SEP-2000.
PF	XX	
XX	XX	25-FEB-2000; 2000EP-0301439.
XX	XX	
XX	XX	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999;	99US-01231801;
PR	09-MAR-1999;	99US-01235481;
PR	23-MAR-1999;	99US-01257681;
PR	25-MAR-1999;	99US-01262641;
PR	29-MAR-1999;	99US-01267851;
PR	01-APR-1999;	99US-01274621;
PR	06-APR-1999;	99US-01282341;
PR	08-APR-1999;	99US-01287141;
PR	16-APR-1999;	99US-01298451;

PR	26-OCT-1999;	99US-01613559;
PR	26-OCT-1999;	99US-01613601;
PR	26-OCT-1999;	99US-01613611;
PR	28-OCT-1999;	99US-01619201;
PR	28-OCT-1999;	99US-01619202;
PR	28-OCT-1999;	99US-01619931;
PR	28-OCT-1999;	99US-01621421;
PR	29-OCT-1999;	99US-01621422;
Query Match		
Best Local Similarity		39.6%;
Matches 242;		Conservative 62.2%;
QY	226	ELFVPEKLNKLDHVRRAEATL
DB	60	ELIVEPKRREKKHEAADDL
QY	286	L-LDDG-VTNMSIPVLVDV
DB	120	LRLLDGSVYNMSVPVLAI
QY	343	SRVYGTCTKKPHI-KMYVM
DB	180	ARTWGTATGLPYVDEAIT
QY	402	ADAVFAFQLRNPVNHGHAL
DB	240	ADAVFAFQLRNPVNHGHAL
QY	462	HAALVEEGVLDPKSTIVAI
DB	300	HEKVLVDGVLDPETTVISL
QY	522	TKKLDYEPHGHGKVLSPAP
DB	360	EKRLDYADHGKVKVLSMAP
QY	582	RKLAREGNPPDGFMAPKAK
DB	420	RTLAKNNENPPDGFMCPCGG
RESULT 9		
AAG30254		
ID	AAG30254	standard; Protein;
XX	XX	AAG30254;
AC	AC	
DT	DT	17-OCT-2000 (first entry)
XX	XX	Arabidopsis thaliana protein
DE	DE	Protein identification; signifi-
KW	KW	hybridisation assay; genetic
KW	KW	termination sequence.
XX	XX	Arabidopsis thaliana.
XX	XX	EP1033405-A2.
XX	XX	06-SEP-2000.
PD	PD	
XX	XX	25-FEB-2000; 2000EP-0301439.
XX	XX	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999;	99US-0123180.
PR	09-MAR-1999;	99US-0123548.
PR	23-MAR-1999;	99US-0125768.
PR	25-MAR-1999;	99US-0126264.
PR	29-MAR-1999;	99US-0126785.
PR	01-APR-1999;	99US-0127462.
PR	06-APR-1999;	99US-0128234.
PR	08-APR-1999;	99US-0128714.
PR	16-APR-1999;	99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-APR-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

[illegible]

RESULT 11
AAG30252
ID AAG30252 standard; Protein; 510 AA.
XX AC AAG30252;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 36134.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141642.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR	20-AUG-1999;	99US-0149722.	
PR	20-AUG-1999;	99US-0149723.	
PR	20-AUG-1999;	99US-0149929.	
PR	23-AUG-1999;	99US-0149902.	
PR	23-AUG-1999;	99US-0149930.	
PR	25-AUG-1999;	99US-0150566.	
PR	26-AUG-1999;	99US-0150884.	
PR	27-AUG-1999;	99US-0151085.	
PR	27-AUG-1999;	99US-0151086.	
PR	27-AUG-1999;	99US-0151080.	
PR	30-AUG-1999;	99US-0151303.	
PR	31-AUG-1999;	99US-0151348.	
PR	01-SEP-1999;	99US-0151930.	
PR	07-SEP-1999;	99US-0152363.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160960.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	

Query Match

Best Local Similarity 39.2%; Score 1282; DB 21; Length 510;

Mismatches 60; Mismatches 81; Indels 6; Gaps 5;

QY	226	ELFVPEPNKLDHVAEATLPSLSITKLDLOWOVLSEGNATPLKGFMRKEYLQVMHFD	285
Db	107	DLVPEPRRREKHEADLPVRVLTATDLQMHVLSGNWASPLRGFMRESEFLQTLHNL	166
QY	286	L-LDDG-VINMSPTIVLPSAEDKTRLEGCSKV--LAHGRRVAILRDAEFYEHKKEER	341
Db	167	LNDDGVSVMSPVIVLADDOQA-LIGESKRVSLVSDSDNPAILNDIEYKHPKEER	225
QY	342	CSRWGTTCTKHPHI-KVMNESGDLVGGDLQVLEKIRWNGLDQVRLTPLELKQCKEM	400

Db	226	IARTWGTAPGLPVVEEAITNAGDWLIGDLEVLVPVKYNDGLDRFLRSPFELKELEKR	285
QY	401	NADAVFAFQLRNPVHNGHALLMQDTCRRLLRGYKHPVLLHPLGGWTKDDVPLDWRMK	460
Db	286	GADAVFAFQLRNPVHNGHALLMTDTRRLLEMGYKNPILLHPLGGFTKADDVPLSWRMK	345
QY	461	OHAALVEEGVLDPKSTIVAFPPSPMLYAGPTEVQWCHRSRMAGANFYIVGRDPAGMHP	520
Db	346	QHEKVLGDGVDPETTVVSIFPPSPMLYAGPTEVQWHAHAKARINAGANFYIVGRDPAGMHP	405
QY	521	ETKKDLYEPHTGGKVLSPAGLTSVEIIPRVAAYNKKAKMDFYDPAHNEFDIFSSTR	580
Db	406	VERDLYADHGGKVLSPAGLERNLIPRVAAYDKTQGMKMAFFDPSRAQDFLIFSGTK	465
QY	581	MRKLAREGENPPDGFMAPKAKVLTIDYRS	610
Db	466	MRALAKNRENPPDGFMCPPGGMKVLVDYDS	495
RESULT 12			
AAG09595			
ID	AAG09595 standard; Protein; 473 AA.		
XX	AC AAG09595;		
XX	17-OCT-2000 (first entry)		
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 7590.		
DE	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
KW	termination sequence.		
XX	Arabidopsis thaliana.		
XX	EP1033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-0301439.		
XX	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126264.		
PR	29-MAR-1999; 99US-0126785.		
PR	01-APR-1999; 99US-0127462.		
PR	06-APR-1999; 99US-0128234.		
PR	08-APR-1999; 99US-0128714.		
PR	16-APR-1999; 99US-0129845.		
PR	19-APR-1999; 99US-0130077.		
PR	21-APR-1999; 99US-0130449.		
PR	23-APR-1999; 99US-0130510.		
PR	28-APR-1999; 99US-0130891.		
PR	30-APR-1999; 99US-0131449.		
PR	04-MAY-1999; 99US-0132048.		
PR	05-MAY-1999; 99US-0132484.		
PR	06-MAY-1999; 99US-0132486.		
PR	06-MAY-1999; 99US-0132487.		
PR	11-MAY-1999; 99US-0132863.		
PR	14-MAY-1999; 99US-0134218.		
PR	14-MAY-1999; 99US-0134219.		
PR	14-MAY-1999; 99US-0134221.		
PR	14-MAY-1999; 99US-0134370.		
PR	18-MAY-1999; 99US-0134768.		
PR	19-MAY-1999; 99US-0134941.		
PR	20-MAY-1999; 99US-0135124.		
PR	21-MAY-1999; 99US-0135353.		
PR	24-MAY-1999; 99US-0135629.		

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.

```
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      38.9%; Score 1272.5; DB 21; Length 473;
Best Local Similarity 57.0%; Pred. No. 1.4e-119;
Matches 243; Conservative 72; Mismatches 92; Indels 19; Gaps 6;

QY 207 VELLOQNIVPYTIK-----DIHELFPENKLDHVRAEATLPSLSTIKD 253
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 39 LNLVYKRNLTMOSVSKMTVKSSLLDPDGGELVELIVPETEIGVKKRAESETMPKVLNQID 98
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 254 LQWQVLSGNATPLKGFNREYQLQVHFDL-LDDGV-INMSIPVLPVSAEDKTRLE 311
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 99 LEWVHVISEGWSPLKGFNREDEYQLQSLHFNLSRLKNGTFVNMSPVLVLAIDDDTKQET- 157
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 312 GCSKFV--LAHGGRVAILRDAAEFYEHKREERCSRWGTTCTKHPI-KMVMESGDWLG 368
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 158 GSENVALVCPGGDIIIGSLSRSEIYKHKEERARTWGTSPGLPVEEYITPSGNWLG 217
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 369 GDQVLEKIRWNGLDQYRLTLELKOKCKEMNADAVAFQLRPNVHNGHALLMDQTCRR 428
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 218 GDLVEFEPKYNDGLDHYRLSPKQREEDNRQADAVPAFQLRPNVHNGHALLMDTRKR 277
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 429 LLERGYKHPVLLHPLGGTCKDDVPLDWRMKQAAVLEEGVLDPKSTIVAFPSPLMYA 488
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 278 LLEMGYKNVLLHPLGGTCKADDVPLDVRMEQHSKVLDDGVLDPKTTIVSFPSPMHYA 337
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 489 GTEVQWHCRSMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLMSAPGLTSVEII 548
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 338 GTEVQWHAKARINAGANFYIVGRDPAGMGHPTEKRDLYDPDQKRVLSMAFGLEKLNIL 397
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 549 PFRVAAYNKAAMFYDYPARINEFDIFSGTRMKRLAREGENPDGCFMAPKAWKVLTQY 608
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 398 PFRVAAYDIEKKMAFFDPSRAKEFLISGTRMRTYARTGENPDGFCMPSGWNVLKYY 457
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 609 RSEMDK 614
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 458 ESLQEK 463
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
AAG09594
ID AAG09594 standard; Protein; 476 AA.
XX
AC AAG09594;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 7589.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132486.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
```

```

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145102.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145911.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

```

```

PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 38.9%; Score 1272.5; DB 21; Length 476;
Best Local Similarity 57.0%; Pred. No. 1.4e-119;
Matches 243; Conservative 72; Mismatches 92; Indels 19; Gaps 6;

Qy 207 VELLQEQNIVPYTIK-----DIHFLFVPEKNKLDHVRARAEITLPSLSITKLD 253
Db 42 LNLVYKRNLTQSVSKMTVTKSSLLDDPGGELVELIVPTEIGVKKAESETMPKVKLLNOLD 101

Qy 254 LOWQVILSEGWATPLKGFMRKEKEYLQVMHFDL-LDDGV-INMSIPVLPVSAEDKTRLE 311
Db 102 LEWVHVISEGWASPLKGFMRDEYLSLHFNLSRLKNGTFVMSLPIVLAIDDDTKEQI- 160

Qy 312 GCSKEV--LAHGRRVAILRDAEYEHKRECRSVGTCTTKPHI-KVMESGDWLVG 368
Db 161 GSSENVALVCPQGGDIIGLSRSVETVYKHKEERARTGTGTTSPGLPYVEEYITPSCNWLIG 220

Qy 369 GDLOVLEKIRWNGDLOYRLTPLELKOKCKEMNADAVAFOLRNPVHNGHALLMODTCRR 428
Db 221 GDLEVFEPITKNDGLDHYRUSPKOLREEFNRQADAVAFOLRNPVHNGHALLMNDTKR 280

Qy 429 LLERGYKHPVLLHPLGGWTKDDVPLDWRMKQHAALVEEGVLDPKSTIVAIFPSPMLYA 488
Db 281 LLENGYKNPVLLHPLGGFTKADDVPLDVRMEQHSKVLDEGVLDPKTTIVSIFSPWHYA 340

Qy 489 GPTEVQWCHRSRMITAGANFYIVGRDPAGMPHPETKKDLYEPHGGKVLSMAPGLTSVEII 548
Db 341 GPTEVQWHAARINAGANFYIVGRDPAGMGHPTEKRDLYDPDHGKRVLSMAPGLEKNIL 400

Qy 549 PFRVAAYNKAKAMDYDPAHNEFDFTSGTRMKLAREGENPPDGFMAPKAWKVLTDYY 608
Db 401 PFRVAAYDTIEKNMAFFDPSRAKEFLFTSGTKMTYARTGENPPDGFMCPSGWNVLKYY 460

Qy 609 RSEMDK 614
Db 461 ESLOEK 466

```

RESULT 14

AAG09596
ID AAG09596 standard; Protein; 425 AA.

XX AC AAG09596;

XX DT 17-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 7591.

XX KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter;

```
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 29-MAR-1999; 99US-0126264.
XX 01-APR-1999; 99US-0126785.
XX 06-APR-1999; 99US-0127462.
XX 08-APR-1999; 99US-0128234.
XX 16-APR-1999; 99US-0128714.
XX 19-APR-1999; 99US-0129845.
XX 21-APR-1999; 99US-0130077.
XX 23-APR-1999; 99US-0130449.
XX 28-APR-1999; 99US-0130510.
XX 30-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147933.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
```

PR	10-SEP-1999;	99US-0153070.	
PR	13-SEP-1999;	99US-0153758.	
PR	15-SEP-1999;	99US-0154018.	
PR	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160989.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161922.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 38.8%; Score 1270; DB 21; Length 425;			
Best Local Similarity 60.5%; Pred. No. 2.1e-119;			
Matches 240; Conservative 65; Mismatches 86; Indels 6; Gaps 5;			
QY	223	DIHELFPENKLDHVRAEPLPSLSITKLDLQWQVLSSEGWATPLKGFMRKEYLQVMH	282
Db	20	ELVELIVPETETGVKKAESMPKVKLNQIDLEWHVHVSSEGWASPLKGFMRDEYLSLH	79
QY	283	FDTL-LDGV-INMSIPVLPSAEDKTRLEGCSKFV--LAHGGRRVAILRDAEFYHRK	338
Db	80	FNSLRKNGTFVNSLPVLVLDLDDTKQOI-GSENVALVCPQDIIIGSLRSVEIYKRNK	138
QY	339	EECRSVKGTCTTKPHI-KVMESGDWLVGDLQVLEKIRWNGDQYRLTLELKOC	397
Db	139	EERITARTGTTSPGLPYVEEYITPSGNWLGDLVFEPIKYNQDGLHYRLSPKQLREF	198
QY	398	KEMNADAVATQLRNPVINGHALLMODTCRLLERGYKHPVLLHPLGGTQKDDVPDLDW	457
Db	199	DNKQADAVAFQLRNPVINGHALLMNDTRKRLLENGYKPNVLLHPLGGTQKDDVPDLDV	258
QY	458	RMKOHAAVLESGDLPKPTIVAIPTSPMLYAGTPEVQHCRSMIAGANFYIVGRDPAGM	517
Db	259	RMEQHSKVLGDLVDPKPTIIVSIPSPHYAGTPEVQWAKARINAGANFYIVGRDPAGM	318
QY	518	PHPEFKDLYEPTHGGKVLSPAGLTSVEIIPFRVAAYNKAAMDFDYPARHNEFDIS	577
Db	319	GHPTKRDLYDPDHGKRVLSNAPGLEKNILPFRVAAYDTIEKKWAFDPSAKEFLIS	378
QY	578	GTRMKKLAREGENPPDGFMAPKAWKVLTDIYRSEMDK	614
Db	379	GTKMRTYARTGENPPDGFMCPSGWNVLVKYIESLOEK	415
RESULT 15			
AAG43667			
ID	AAG43667 standard; Protein; 408 AA.		
XX	AAG43667;		
XX	18-OCT-2000 (first entry)		
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 54606.		
XX	Protein identification; signal transduction pathway; metabolic pathway;		
XX	hybridisation assay; genetic mapping; gene expression control; promoter;		
XX	termination sequence.		
OS	Arabidopsis thaliana.		
PN	EP1033405-A2.		
XX	06-SEP-2000.		
XX	25-FEB-2000; 2000EP-0301439.		
XX	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126264.		
PR	29-MAR-1999; 99US-0126785.		
PR	01-APR-1999; 99US-0127462.		
PR	06-APR-1999; 99US-0128234.		
PR	08-APR-1999; 99US-0128714.		
PR	16-APR-1999; 99US-0129845.		
PR	19-APR-1999; 99US-0130077.		
PR	21-APR-1999; 99US-0130449.		
PR	23-APR-1999; 99US-0130510.		
PR	28-APR-1999; 99US-0130891.		
PR	30-APR-1999; 99US-0131449.		
PR	30-APR-1999; 99US-0132048.		
PR	04-MAY-1999; 99US-0132407.		
PR	05-MAY-1999; 99US-0132484.		
PR	06-MAY-1999; 99US-0132485.		
PR	06-MAY-1999; 99US-0132486.		
PR	07-MAY-1999; 99US-0132487.		
PR	11-MAY-1999; 99US-0132863.		
PR	14-MAY-1999; 99US-0134256.		
PR	14-MAY-1999; 99US-0134218.		
PR	14-MAY-1999; 99US-0134219.		
PR	14-MAY-1999; 99US-0134221.		
PR	14-MAY-1999; 99US-0134370.		
PR	18-MAY-1999; 99US-0134768.		
PR	19-MAY-1999; 99US-0134941.		
PR	20-MAY-1999; 99US-0135124.		
PR	21-MAY-1999; 99US-0135353.		
PR	24-MAY-1999; 99US-0135629.		
PR	25-MAY-1999; 99US-0136021.		
PR	27-MAY-1999; 99US-0136392.		
PR	28-MAY-1999; 99US-0136782.		
PR	01-JUN-1999; 99US-0137222.		
PR	03-JUN-1999; 99US-0137528.		
PR	04-JUN-1999; 99US-0137502.		
PR	07-JUN-1999; 99US-0137724.		
PR	08-JUN-1999; 99US-0138094.		
PR	10-JUN-1999; 99US-0138540.		
PR	10-JUN-1999; 99US-0138847.		
PR	14-JUN-1999; 99US-0139119.		
PR	16-JUN-1999; 99US-0139452.		
PR	17-JUN-1999; 99US-0139453.		
PR	17-JUN-1999; 99US-0139492.		

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144633.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 38.5%; Score 1259.5; DB 21; Length 408;
Best Local Similarity 60.7%; Pred. No. 2.3e-118;
Matches 236; Conservative 69; Mismatches 79; Indels 5; Gaps 5;

Qy 227 LFVPENKLDHVRAEATLPS-LSITKLDLOWOVLSGWTPLKGFMEKEEYLQVWHD 285
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 3 LVVEESRRVRWKHEATVPARIKLNRLVDLEWHLVLSGWNASPLKGFMEKEEYLQVWHD 62
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Qy 286 L-LDDG-VINMISIPVLPVSAEDKTRLEGCSEKFLAHG-GRRVAILRDAEFYEHKREERC 342
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
Db 63 FRLEDGVSVMNSVPVILADDDQKFRIGDSNQVTLVDVSGNPIALNDIEYKHPKEERI 122

Search completed: May 30, 2003, 12:17:32
Job time : 76 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:16:34 ; Search time 27 Seconds
(without alignments)
670.189 Million cell updates/sec

Title: US-09-898-165B-7

Perfect score: 3271

Sequence: 1 MSICIKKQKTEKQKSTNVVY.....MAPKAWKVLTDYRSEMDKN 615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No...is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2602	79.5	624	2	US-08-879-561-3
2	2599	79.5	624	2	US-08-879-561-9
3	2367	72.4	610	2	US-08-879-561-10
4	1274	38.9	463	4	US-09-346-408-11
5	1268.5	38.8	476	4	US-09-346-408-12
6	1263	38.6	465	4	US-09-346-408-6
7	1237.5	37.8	465	4	US-09-346-408-4
8	1227	37.5	461	4	US-09-346-408-8
9	562	17.2	133	4	US-09-346-408-2
10	481	14.7	204	4	US-09-134-001C-4517
11	442.5	13.5	401	4	US-09-134-001C-4403
12	417	12.7	110	4	US-09-346-408-10
13	323	9.9	174	4	US-09-149-476-635
14	198	6.1	36	4	US-09-149-476-634
15	106	3.2	816	1	US-08-038-760-3
16	106	3.2	816	2	US-08-470-091-3
17	106	3.2	928	1	US-08-204-329-1
18	106	3.2	928	2	US-08-959-638-8
19	106	3.2	928	2	US-08-482-627-5
20	106	3.2	928	3	US-08-801-092-4
21	106	3.2	928	4	US-08-328-673A-8
22	106	3.2	928	4	US-09-315-113-4
23	106	3.2	928	5	PCT-US94-10357-2
24	104.5	3.2	5588	4	US-09-036-987A-6
25	104.5	3.2	5588	4	US-09-370-700-6
26	104	3.2	928	5	PCT-US94-10357-3
27	98.5	3.0	1365	6	5194600-4

Patent No. 5194600

ALIGNMENTS

RESULT 1

US-08-879-561-3

; Sequence 3, Application US/08879561

; Patent No. 5817482

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Hawkins, Phillip R.

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/879,561

; FILING DATE: Herewith

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0325 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 624 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: LUNGN02

; CLONE: 373887

; US-08-879-561-3

Sequence 5377, Ap
Sequence 29, Appl
Sequence 29, Appl
Sequence 30, Appl
Sequence 30, Appl
Sequence 30, Appl
Sequence 46, Appl
Sequence 18, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 28, Appl
Sequence 28, Appl
Sequence 5, Appl
Sequence 4, Appl
Sequence 119, App
Sequence 42, Appl

Query Match 79.5%; Score 2602; DB 2; Length 624;
Best Local Similarity 77.1%; Pred. No. 6.1e-269;
Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

QY 5 KKOKTEN-----QOKSTNVVYQAHVSRNKRQVVGTRGGFRGCTVWLTGLSGAGKTTI 58
DB 9 KVKLSNNAQNMGMQRATNVTYQAHVSRNKRQVVGTRGGFRGCTVWLTGLSGAGKTTV 68

QY 59 SFALREYLVSHATPCYSLDGDNVVRHGLNRNLGSPGDRENIIRIAEVAKLADAGLVCI 118
DB 69 SNALEEVLYCHGIPCTYLDGDNIRQGLNKLNGSPEDRENVIRIAEVAKLADAGLVCI 128

QY 119 TSFISPAKDRENAKIHESAGLPFFEIFVDAPLNICESRDVKGLYKARAGAIKGTGTGI 178
DB 129 TSFISPYTDORNNARQIHESAGLPFFEIFVDAPLNICESRDVKGLYKARAGAIKGTGTGI 188

QY 179 DSDYEKPEPVRVLTNLSVSCVHQVVELLOEQNIPTVYIIKDHELFVVPENKLDHVR 238
DB 189 DSEYEKPEAPLVKTDSCDVNDVQVQVVELLOEQNIPTVYIIKDHELFVVPENKLDHVR 248

QY 239 AEAETLPSLSITKLDLQWVQVLSSEGWATPLKGFPMREKEYLQVMHFDTLDDGVINNSIPI 298
DB 309 VLTATHEDEKRLDGCCTAFALVYEGRRVAILRNPPEFHEHKEERCARQWGTCKNHPYIKM 368

QY 359 VMEGDMVLGGDLQVLEKIRWNDGLDQYRLTPELQKCKEMNADAVFAFOLRNPVHNGH 418
DB 369 VMEQDMLTGGDLQVLDRIYVNDGLDQYRLTPELQKCKEMNADAVFAFOLRNPVHNGH 428

QY 419 ALLMQDTCRLLERGLYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLPETTVV 478
DB 429 ALLMQDTHQOLLERYRRPVLHLLPLGGWTKDDVPLDMRMKQHAALVEEGVLPETTVV 488

QY 479 AIFPSPMLYAGPTEVOWHCRSMIAGANFYIYGRDPAGMPHPTKDLPEPHGGKVLMS 538
DB 489 AIFPSPMYAGPTEVOWHCRSMIAGANFYIYGRDPAGMPHPTKDLPEPHGGKVLMS 548

QY 539 APLGTSVEIIPFVAAYNAKAKAMDYDPAHNEFDIFSGTRMKRLAREGENPPDGFMAP 598
DB 549 APLGTLTLEIVPFRVAAYNAKAKAMDYDPAHNEFDIFSGTRMKRLAREGENPPDGFMAP 608

QY 599 KAWKVLTDYRS 610
DB 609 KAWTVLVEYKS 620

RESULT 2

US-08-879-561-9
Sequence 9, Application US/08879561
Patent No. 5817482

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Hawkins, Phillip R.
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.

TITLE OF INVENTION: DISEASE RELATED NUCLEOTIDE KINASES
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/879,561
FILING DATE: Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0325 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 624 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1109876
US-08-879-561-9

Query Match 79.5%; Score 2599; DB 2; Length 624;
Best Local Similarity 77.0%; Pred. No. 1.3e-268;
Matches 471; Conservative 66; Mismatches 69; Indels 6; Gaps 1;

QY 5 KKOKTEN-----QOKSTNVVYQAHVSRNKRQVVGTRGGFRGCTVWLTGLSGAGKTTI 58
DB 9 KVKLSNNAQNMGMQRATNVTYQAHVSRNKRQVVGTRGGFRGCTVWLTGLSGAGKTTV 68

QY 59 SFALREYLVSHATPCYSLDGDNVVRHGLNRNLGSPGDRENIIRIAEVAKLADAGLVCI 118
DB 69 SNALEEVLYCHGIPCTYLDGDNIRQGLNKLNGSPEDRENVIRIAEVAKLADAGLVCI 128

QY 119 TSFISPAKDRENAKIHESAGLPFFEIFVDAPLNICESRDVKGLYKARAGAIKGTGTGI 178
DB 129 TSFISPYTDORNNARQIHESAGLPFFEIFVDAPLNICESRDVKGLYKARAGAIKGTGTGI 188

QY 179 DSDYEKPEPVRVLTNLSVSCVHQVVELLOEQNIPTVYIIKDHELFVVPENKLDHVR 238
DB 189 DSEYEKPEAPLVKTDSCDVNDVQVQVVELLOEQNIPTVYIIKDHELFVVPENKLDHVR 248

QY 239 AEAETLPSLSITKLDLQWVQVLSSEGWATPLKGFPMREKEYLQVMHFDTLDDGVINNSIPI 298
DB 249 TDAEALPALKINKYDMQVQVLAEGWATPLNGFMREREYLOCLHFDCLDGGVINSVPI 308

QY 299 VLPVSAEDKTRLEGCSKFLAHGGRVAILRDAEFYEHKEERCRCRVWGTCTCKHPIKIM 358
DB 309 VLTATHEDEKRLDGCCTAFALVYEGRRVAILRNPPEFHEHKEERCARQWGTCKNHPYIKM 368

QY 359 VMEGDMVLGGDLQVLEKIRWNDGLDQYRLTPELQKCKEMNADAVFAFOLRNPVHNGH 418
DB 369 VMEQDMLTGGDLQVLDRIYVNDGLDQYRLTPELQKCKEMNADAVFAFOLRNPVHNGH 428

QY 419 ALLMQDTCRLLERGLYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLPETTVV 478
DB 429 ALLMQDTHQOLLERYRRPVLHLLPLGGWTKDDVPLDMRMKQHAALVEEGVLPETTVV 488

QY 479 AIFPSPMLYAGPTEVOWHCRSMIAGANFYIYGRDPAGMPHPTKDLPEPHGGKVLMS 538
DB 489 AIFPSPMYAGPTEVOWHCRSMIAGANFYIYGRDPAGMPHPTKDLPEPHGGKVLMS 548

QY 539 APLGTSVEIIPFVAAYNAKAKAMDYDPAHNEFDIFSGTRMKRLAREGENPPDGFMAP 598
DB 549 APLGTLTLEIVPFRVAAYNAKAKAMDYDPAHNEFDIFSGTRMKRLAREGENPPDGFMAP 608

QY 599 KAWKVLTDYRS 610
DB 609 KAWTVLVEYKS 620


```
Db 300 HEWLEDCVLDPETTVVSIIPSPMHYACPTVEQVHAKARINAGANFIVIGRDPAGMGHPL 359
Qy 522 TKDLYEPTHGKVL$MAPGLTSVEIIPFRVAAYNKAAMDFYDPAHNEFFDISGTRM 581
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 360 EKRDLYDADHGKVL$MAPGLERLNLIPKVAAYDKTQNGMAFFDPSRPQDFLISGTM 419
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 582 RKLAREGENPPDGFMAPKANKVLTDDYRS 610
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 420 RALAKNKENPPDGMFCGGKVLVDYDYS 448

RESULT 5
US-09-346-408-12
; Sequence 12, Application US/09346408B
; Patent No. 6338966
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falco, Carl
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-A
; CURRENT APPLICATION NUMBER: US/09/346.408B
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 60/092.833
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-346-408-12

Query Match 38.8%; Score 1268.5; DB 4; Length 476;
Best Local Similarity 57.3%; Pred. No. 1.3e-126;
Matches 242; Conservative 71; Mismatches 90; Indels 19; Gaps 6;

Qy 207 VELLOEGNIVPYIIRK-----DIHELFPENKLDHVRARAEETLPSLSTKLD 253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 42 LNLVYRNLTMQSVKMTVKSSLDPDGGELVELIIPETEIGVKKAESETMPKVKLNQID 101
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 254 LQWQVLSSEGWA$PLKGFMEKEVYLOVMHFDTL-LDDGV-INMSIPVLVPSAEDKTRLE 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102 LEWVHVISGWA$PLKGFMEDEYLOSLSHNSRLKNGTFVNSLPIVLAI$DDTKEQI- 160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 312 GCSKFV--LAHGRRVAILRDAEFYEHKREERC$RVWGTCTTKPHI-KMVMSGDLVYG 368
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 161 GSSENVALVCPQDIIIGSLRSVETIKHKEERARTWGTSPGLPYVEEYITPSGNWLI 220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 369 GDLOVLEKIRWNGLDQYRLTPLELKOKCKEMNADAVAFQLRNPVHNGHALLMQDTCRR 428
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221 GDLEVFETIKYNDGLDHYRLSPKQLEEFNROQADAVAFQLRNPVHNGHALLMNDTRK 280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 429 LLERGKHPVLLHPLGGMWKDDVPLDWRMKOHA$VLEBGLDPKSTIVAIPSPMLYA 488
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 281 LLEMGYNPVLHPLGFGTKADDVPLDVRMEGHSKVL$EDGVLDPKTTIVSIIPSPMHYA 340
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 489 GPTEVQWHCSRMIAGANFYIVGRDPAGMPHPETKKDLYEP$THGGKVL$MAPGLTSVEII 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 341 GPTEVQWHAKARINAGANFYIVGRDPAGMGHPTEKRDLYDPDHGKVL$MAPGLEKLNIL 400
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 549 PFRVAAYNKAAMDFYDPAHNEFFDISGTRMRKLAREGENPPDGFMAPKANKVLTDDY 608
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 401 PFRVAAYDTIEKKMAFFDPSRAKEFLFISGTRMRTYARTGENPPDGMFCPSGNWLVKY 460
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 609 RS 610
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 461 ES 462

RESULT 6
US-09-346-408-6
```

```
; Sequence 6, Application US/09346408B
; Patent No. 6338966
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falco, Carl
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-A
; CURRENT APPLICATION NUMBER: US/09/346.408B
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 60/092.833
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Glycine max
US-09-346-408-6

Query Match 38.6%; Score 1263; DB 4; Length 465;
Best Local Similarity 61.7%; Pred. No. 4.8e-126;
Matches 240; Conservative 59; Mismatches 86; Indels 4; Gaps 4;

Qy 226 ELFVPEKNKLDHVRARAEETLPSLSTKLDLQWQVLSSEGWA$PLKGFMEKEVYLOVMHFD 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 ELVVTDFERDLKKG$ALSRLPRIKSLRIDLEWHVLSSEGWA$PLKGFMEKEVYLOVTHFNS 121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 286 L-LDDG-VINMSIPVLVPSAEDKTRLEGC$KFLAHG-GRRVAILRDAEFYEHKREERC 342
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 LRLDGSVVMNSVPIVLAI$DDAQKHRI$GDNK$KVALFDSK$GDPVAILNII$YKPKKEERI 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 343 SRVWGTCTTKPHIKM-VMSGDLVGGDLQVLEKIRWNGLDQYRLTPLELKOKCKEMN 401
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 182 ARTWGTIAPGLTYVEQ$ITWAGNWLIG$DLEVEPIQYNDGLDHFRL$PAQLRAEFTRRN 241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 402 ADAVAFQLRNPVHNGHALLMQDTCRLLERGKHPVLLHPLGGMWKDDVPLDWRMKQ 461
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 ADAVAFQLRNPVHNGHALLMTDTRKLEMGYNPVLHPLGGMWKDDVPLDWRMKQ 301
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 462 HAAVLEBGLDPKSTIVAIPSPMLYAGPTEVQWHCSRMIAGANFYIVGRDPAGMPHP 521
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 302 HEVLEDCVLDPETTVVSIIPSPMHYACPTVEQVHAKARINAGANFYIVGRDPAGMHPV 361
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 522 TKDLYEPTHGKVL$MAPGLTSVEIIPFRVAAYNKAAMDFYDPAHNEFFDISGTRM 581
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 362 EKRDLYDADHGKVL$MAPGLERLNLIPKVAAYDKTQNGMAFFDPSRPQDFLISGTM 421
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 582 RKLAREGENPPDGFMAPKANKVLTDDYRS 610
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 422 RTLARNKESPPDGMFCGGKVLVDYDYS 450

RESULT 7
US-09-346-408-4
; Sequence 4, Application US/09346408B
; Patent No. 6338966
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falco, Carl
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-A
; CURRENT APPLICATION NUMBER: US/09/346.408B
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 60/092.833
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 465
```

226 ELFPENKLDHVRAEATLSLSSTKLDLQWQVLSGAWATPLKGFMRREYLLQVMHFD 285
; Patent NO. 0500370
; GENERAL INFORMATION:

APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIORITY FILING DATE: 1997-11-08
PRIORITY FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4517
LENGTH: 204
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4517

Query Match 14.7%; Score 481; DB 4; Length 204;
Best Local Similarity 44.5%; Pred. No. 4.5e-43;
Matches 93; Conservative 37; Mismatches 73; Indels 6; Gaps 2;

QY 8 KTNQOKSTNVVQAHVSRNKRQVVGTRGGRGCTVWLTGLSGAGKTTISFALBEYLV 67
DB 2 RSVKMSBNHITWHDSEVTKQRQ-----KNGHKSAVIWTGLSGSKSTVSVALEKELF 57
QY 68 SHAIPCYSLDGNVRHGLNRLGFSFGDRENRIRRAEVAKLADAGLVCITSFISPPFAK 127
DB 58 NEGQTYRLDGDNRHGLNRLGFSFGDRENRIRRAEVAKLADAGLVCITSFISPPFAK 117
QY 128 DRENNARKIHESAGLPFFPEIVDAPLNICESRDYKGLYKRRAGEIKGFTGIDSDYKPEP 167
DB 118 DREGVRLLEDN--EFIEVYTKSVCEKRDPKGLYKRRAGEIKGFTGIDSDYKPEP 175
QY 188 PERVLKTNLSTVSDCVHQVVELLOEQNV 216
DB 176 PETITDTEHTDIEQSVVQIIRYLKEHEY 204

RESULT 11

US-09-134-001C-4403
Sequence 4403, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIORITY FILING DATE: 1997-11-08
PRIORITY FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4403
LENGTH: 401
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4403

Query Match 13.5%; Score 442.5; DB 4; Length 401;
Best Local Similarity 29.9%; Pred. No. 1.9e-38;
Matches 122; Conservative 82; Mismatches 175; Indels 29; Gaps 11;

QY 206 VVELLOEQNVITYIKDITHELF---VPENKLDHVAEAEITLPSLITKLDLQWQVLSL 262
DB 7 LIIMSNNETITNTYIKPHGGLINRVVEGNERLIEEALKFKPITLNPWGISDLELIGI 66
QY 263 GNATPLKGFMRKEYLQVNHFDITLDDGVNNSIPVLVPSAEKTRLE-GCSKFFVLAHG 321
DB 67 GGFSPUTGFMNKEDYTKVIE-ETHLSNGLV-WSIPITLPTVTESEADKLEIGDIALYGBD 124
QY 322 GRRVAILRDAEYFHRKEKRCRSRVGTCTCKPHKIKMVMESGDLVGGDLQVLEKIRWD 381

DB 125 GQLYGTILKEEYTYDKREARLVYGTTEBAHPGVKVEKGNILGPIKLLNRPK-HD 183
QY 382 GLDOYRLTLPLEKOKCKEMNADAVAFQALRNPVHNGHALLMODTCRLLRGGYKHPVLLL 441
DB 184 AFSNYHLDPSRQLFHDHGLGKNTVVGQTRNPVHRAHEYI-QKSALDIVDG-----LLL 236
QY 442 HPLGGWTKDDVPLDMRMKQAAVLEGVLDPKASTIVAIFPSPMLYAGPTEVQWHCGRM 501
DB 237 NPLVGETKSDIPADVRMESYEVLKKNYPEDRARLV-IYPAAMRYAGPREAILHATVRK 295
QY 502 IAGANFYIVGRDPAGMPHPETKDDLYEPHGGKVLSSMAPGLTSVEIIPFRVAAY-----N 556
DB 296 NYGCTHFIIVGRDHAGV-----GDYGTVEAQLITOFDELDGQILKFEHAFYCEACGN 349
QY 557 KAKKAMDYDPARHNEFDFISGTRMRKLAREGNPPDGENAPKAWKVL 604
DB 350 MATAKTCPHDASQHLH---LSGKVKREKLNGESLPTKFSRPEVAEVL 394

RESULT 12

US-09-346-408-10
Sequence 10, Application US/09346408B
Patent No. 6338966
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Anderson, Shawn
APPLICANT: Falco, Carl
APPLICANT: Rafalski, Antoni
TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
FILE REFERENCE: BB-1167-A
CURRENT APPLICATION NUMBER: US/09/346,408B
CURRENT FILING DATE: 1999-07-01
EARLIER APPLICATION NUMBER: 60/092,833
EARLIER FILING DATE: July 14, 1998
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Microsoft Office 97
SEQ ID NO 10
LENGTH: 110
TYPE: PRT
ORGANISM: Triticum aestivum
US-09-346-408-10

Query Match 12.7%; Score 417; DB 4; Length 110;
Best Local Similarity 71.3%; Pred. No. 1.1e-36;
Matches 77; Conservative 14; Mismatches 17; Indels 0; Gaps 0;

QY 359 VMESGDLVGGDLQVLEKIRWNGDQYRLTLPLEKOKCKEMNADAVAFQALRNPVHNGH 418
DB 3 ITNAGDLVGGDLQVLEKIRWNGDQYRLTLPLEKOKCKEMNADAVAFQALRNPVHNGH 62
QY 419 ALLMODTCRLLRGGYKHPVLLLHPLGGWTKDDVPLDMRMKQAAV 466
DB 63 ALLMTDTRRLLEMGYKHPVLLLHPLGGWTKDDVPLDMRMKQAAV 110

RESULT 13

US-09-149-476-635
Sequence 635, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002P1
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621

1	EARLIER FILING DATE: 1997-03-07	2
3	EARLIER APPLICATION NUMBER: 60/040,626	4
5	EARLIER FILING DATE: 1997-03-07	6
7	EARLIER APPLICATION NUMBER: 60/040,334	8
9	EARLIER FILING DATE: 1997-03-07	10
11	EARLIER APPLICATION NUMBER: 60/040,336	12
13	EARLIER FILING DATE: 1997-03-07	14
15	EARLIER APPLICATION NUMBER: 60/040,163	16
17	EARLIER FILING DATE: 1997-03-07	18
19	EARLIER APPLICATION NUMBER: 60/047,600	20
21	EARLIER FILING DATE: 1997-05-23	22
23	EARLIER APPLICATION NUMBER: 60/047,615	24
25	EARLIER FILING DATE: 1997-05-23	26
27	EARLIER APPLICATION NUMBER: 60/047,597	28
29	EARLIER FILING DATE: 1997-08-23	30
31	EARLIER APPLICATION NUMBER: 60/047,502	32
33	EARLIER FILING DATE: 1997-05-23	34
35	EARLIER APPLICATION NUMBER: 60/047,633	36
37	EARLIER FILING DATE: 1997-05-23	38
39	EARLIER APPLICATION NUMBER: 60/047,583	40
41	EARLIER FILING DATE: 1997-05-23	42
43	EARLIER APPLICATION NUMBER: 60/047,617	44
45	EARLIER FILING DATE: 1997-05-23	46
47	EARLIER APPLICATION NUMBER: 60/047,618	48
49	EARLIER FILING DATE: 1997-05-23	50
51	EARLIER APPLICATION NUMBER: 60/047,503	52
53	EARLIER FILING DATE: 1997-05-23	54
55	EARLIER APPLICATION NUMBER: 60/047,592	56
57	EARLIER FILING DATE: 1997-05-23	58
59	EARLIER APPLICATION NUMBER: 60/047,581	60
61	EARLIER FILING DATE: 1997-05-23	62
63	EARLIER APPLICATION NUMBER: 60/047,584	64
65	EARLIER FILING DATE: 1997-05-23	66
67	EARLIER APPLICATION NUMBER: 60/047,500	68
69	EARLIER FILING DATE: 1997-05-23	70
71	EARLIER APPLICATION NUMBER: 60/047,587	72
73	EARLIER FILING DATE: 1997-05-23	74
75	EARLIER APPLICATION NUMBER: 60/047,492	76
77	EARLIER FILING DATE: 1997-05-23	78
79	EARLIER APPLICATION NUMBER: 60/047,598	80
81	EARLIER FILING DATE: 1997-05-23	82
83	EARLIER APPLICATION NUMBER: 60/047,613	84
85	EARLIER FILING DATE: 1997-08-23	86
87	EARLIER APPLICATION NUMBER: 60/047,582	88
89	EARLIER FILING DATE: 1997-05-23	90
91	EARLIER APPLICATION NUMBER: 60/047,596	92
93	EARLIER FILING DATE: 1997-05-23	94
95	EARLIER APPLICATION NUMBER: 60/047,612	96
97	EARLIER FILING DATE: 1997-05-23	98
99	EARLIER APPLICATION NUMBER: 60/047,632	100
101	EARLIER FILING DATE: 1997-05-23	102
103	EARLIER APPLICATION NUMBER: 60/047,601	104
105	EARLIER FILING DATE: 1997-05-23	106
107	EARLIER APPLICATION NUMBER: 60/043,580	108
109	EARLIER FILING DATE: 1997-04-11	110
111	EARLIER APPLICATION NUMBER: 60/043,568	112
113	EARLIER FILING DATE: 1997-04-11	114
115	EARLIER APPLICATION NUMBER: 60/043,314	116
117	EARLIER FILING DATE: 1997-04-11	118
119	EARLIER APPLICATION NUMBER: 60/043,569	120
121	EARLIER FILING DATE: 1997-04-11	122
123	EARLIER APPLICATION NUMBER: 60/043,311	124
125	EARLIER FILING DATE: 1997-04-11	126
127	EARLIER APPLICATION NUMBER: 60/043,671	128
129	EARLIER FILING DATE: 1997-04-11	130
131	EARLIER APPLICATION NUMBER: 60/043,674	132
133	EARLIER FILING DATE: 1997-04-11	134
135	EARLIER APPLICATION NUMBER: 60/043,669	136
137	EARLIER FILING DATE: 1997-04-11	138
139	EARLIER APPLICATION NUMBER: 60/043,312	140
141	EARLIER FILING DATE: 1997-04-11	142
143	EARLIER APPLICATION NUMBER: 60/043,313	144
145	EARLIER FILING DATE: 1997-04-11	146

EARLIER APPLICATION NUMBER: 60/043,672	EARLIER FILING DATE: 1997-04-11	EARLIER APPLICATION NUMBER: 60/043,315	EARLIER FILING DATE: 1997-04-11	EARLIER APPLICATION NUMBER: 60/048,974	EARLIER FILING DATE: 1997-06-06	EARLIER APPLICATION NUMBER: 60/056,886	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,877	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,889	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,893	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,630	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,878	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,666	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,872	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,882	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,637	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,903	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,888	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,879	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,880	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,894	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,910	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,864	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,631	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,845	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/056,892	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/057,761	EARLIER FILING DATE: 1997-08-22	EARLIER APPLICATION NUMBER: 60/047,595	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,599	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,588	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,585	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,586	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,590	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,594	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,589	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,593	EARLIER FILING DATE: 1997-05-23	EARLIER APPLICATION NUMBER: 60/047,614	EARLIER FILING DATE: 1997-05-23
--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------	--	---------------------------------

```

; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

```

Query Match 9.98; Score 323; DB 4; Length 174;

Best Local Similarity 83.38; Pred. No. 2,6e-26;

Matches 65; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 MSGIKKOKTENQOKSTNVVQAHVSRNKRQGVVTRGGRGCTVWLTGLSGAGKTTISF 60

Db 55 MSGIKKOKTENQOKSTNVVQAHVSRNKRQGVVTRGGRGCTVWLTGLSGAGKNNDKF 114

Qy 61 ALREYLVSHAIPCVSLDG 78

Db 115,CPGVLVSHAIPVNSWNG 132

RESULT 14

US-09-149-476-634

; Sequence 634, Application US/09149476

; Patent No. 6420526

; GENERAL INFORMATION:

; APPLICANT: Rosen-et al.

; TITLE OF INVENTION: 186 Human Secreted proteins

; FILE REFERENCE: P2002p1

; CURRENT APPLICATION NUMBER: US/09/149,476

; EARLIER APPLICATION NUMBER: PCT/US98/04493

; EARLIER FILING DATE: 1998-03-06

; EARLIER APPLICATION NUMBER: 60/040,162

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,333

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/038,621

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,626

; EARLIER FILING DATE: 1997-03-07

```

; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315

```

; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 6.1%; Score 198; DB 4; Length 36;
Best Local Similarity 100.0%; Pred. No. 4e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 YQAHVSRNKRQVVGTRGFRGCTVWLTGLSGACK 55
DB 1 YQAHVSRNKRQVVGTRGFRGCTVWLTGLSGACK 36

RESULT 15
US-08-038-760-3
; Sequence 3, Application US/08038760
; Patent No. 5496731
; GENERAL INFORMATION:
; APPLICANT: Xu, Hong-Ji
; APPLICANT: Hu, Shi-Xue
; APPLICANT: Benedict, William F.
; TITLE OF INVENTION: Broad-Spectrum Tumor Suppressor Genes, Gene Products and
; TITLE OF INVENTION: Methods for Tumor Suppressor Gene Therapy.
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/038,760
; FILING DATE: 19930325
; CLASSIFICATION: 424

Search completed: May 30, 2003, 12:21:02
Job time : 29 secs

ATTORNEY/AGENT INFORMATION:

NAME: Poissant, Brian M
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 7409-025-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-038-760-3

Query Match 3.2%: Score 106; DB 1; Length 816;

Best Local Similarity 18.3%: Pred. No. 0.058;

Matches 139; Conservative 100; Mismatches 224; Indels 296; Gaps 39;

QY 46 WLTGLSGAGKT-----TISFALE-----EYLVSHAIP-----CYSLDGD--NVR 82
DB 83 WITFLAKGEVLQMEDDLVISFQMLCVLDYFIKLSPPMLLKPYKTAVIPINGSPTPR 142
QY 83 HGLNRNLGFSFGDREENIRIAEVAKLFADAGL-----VCITSFISPFADRENARKIHE 137
DB 143 RGNQRSARIA--KOLENDTRIIEVLCKEHCNIDEVKNVYKNEFI-PFM-----NSGLVLT 195
QY 138 SAGLPFEFEIVDAPLNICESRDYKGLYKRARAGEIKGF-----TGIDSDYKPEPT 188
DB 196 SNGLPEVE-----NL--SKRYEEIYKKNKDLDLARFLDHDKTLQTDSDIS-FETQTP 245
QY 189 ERVLKTNLS-----TVSDCVHQVVELLQ-----EQNVIPY-----TI----- 220
DB 246 R--KSNLDEVNVIPPHTPVRTVMNTIQQLMILNSASDQPSLENLISYFNCTVNPKE 302
QY 221 ----IKDIHELFPENKLDHVRAEATLPSLSITKLDLQWVVLSEGWATPLKGFWR--- 273
DB 303 ILKRVKDIGIF---KEKFAKVGQC-----VEIGSORIKLGVRLLYRVME 346
QY 274 -----EKEYLQVMHFDTLDDGVINMSI-----PIVLVPSAEDKTRLEG----- 312
DB 347 SMLKSEERLSIQNFSKLLNDNIFHMSLLACALEVVMATYRSSTSONLSDGTDLSFPWIL 406
QY 313 -----CSKFLVLAHGRVAILRDAEFYHRKEERCVRWGTCTTKHPHKMV 359
DB 407 NVLNLKAFDFYKVIESFIKAEGNLTREMIKHLERCEH-----I 445
QY 360 MESGDWLVGDDL-QVLEKIRWNDG-----LDQYRLTPLELKOK- 396
DB 446 MESLAWLSDSPLDLIKQSKDREGPTDHLASACPLNPLQNNHTAADMY-LSPVRSPPKK 504
QY 397 -----CKEMNADAVAFQLRNP-----VHNGHALLMQDTCRLLERGYKH 436
DB 505 GSTTRVNSTANAETQATSAPOTKPLKSTLSLFYKKVRYLAYRLNLTLCERLS---EH 561
QY 437 PVL-----LLHPLGGWTKDDDVPLDWRM-----KQH 462
DB 562 PELEHIITLQFTLQNEVELMRDRLDQIMMCSMYGICKVKKNIDLKFKIIVTAYKDLPH 621
QY 463 A-----AVLEEGVLDPKSTIVAFIPSPMLYAGPTEVQWHCRSR----- 500
DB 622 AVQETKRVLIKEEYD---SIIVFYNVFMQLKTNILQYASTRPTTSLPIPHIRSPY 678
QY 501 -----MIAGANFYIVGRDPAGMPHETPKKOLYEPHGGKVLMSAPGLTSVEIIPRVA 553
DB 679 KFPSSPLRIPGGNIYI---SPLKSPY-KISEGLPTPT-----KMTPRSRLVSIAGESFG 728
QY 554 AYNKAKAMDFFYPARHNEEDFISCTMRKRLAREGENPP 592
DB 729 TSEKFQKI-----NQMV-CNDSRVLKRSAGSNPP 757

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:19:45 ; Search time 48 Seconds
(without alignments)
1296.929 Million cell updates/sec

Title: US-09-898-165B-7
Perfect score: 3271
Sequence: 1 MSGIKKQKTEKQKSTNVVY.....MAPKAWKLTYYRSEMDKN 615

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 383519 seqs, 101223694 residues
Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_RA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3216	98.3	619	9	US-09-974-298-184
2	2602	79.5	624	9	US-09-984-245-196
3	2602	79.5	624	9	US-09-966-262-196
4	2602	79.5	624	9	US-09-983-966-196
5	2602	79.5	624	9	US-10-143-090-196
6	2588	79.1	624	10	US-09-771-161A-244
7	571.5	17.5	162	9	US-10-156-239-40
8	571.5	17.5	162	9	US-10-199-485-40
9	571.5	17.5	162	10	US-09-795-693-40
10	476.5	14.6	201	10	US-09-741-669-457
11	340	10.4	65	10	US-09-864-761-39981.
12	323	9.9	174	9	US-09-809-391-635
13	221.5	6.8	163	10	US-09-867-550-96
14	217	6.6	41	9	US-09-984-245-340
15	217	6.6	41	9	US-09-966-262-340
16	217	6.6	41	9	US-09-983-966-340
17	217	6.6	41	9	US-10-143-090-340
18	198	6.1	36	9	US-09-809-391-634
19	187	5.7	50	10	US-09-771-161A-153

20	181	5.5	45	9	US-09-984-245-332	Sequence 332, App
21	181	5.5	45	9	US-09-966-262-332	Sequence 332, App
22	181	5.5	45	9	US-09-983-966-332	Sequence 332, App
23	181	5.5	45	9	US-10-143-090-332	Sequence 332, App
24	177	5.4	37	9	US-09-984-245-339	Sequence 339, App
25	177	5.4	37	9	US-09-966-262-339	Sequence 339, App
26	177	5.4	37	9	US-09-983-966-339	Sequence 339, App
27	177	5.4	37	9	US-10-143-090-339	Sequence 339, App
28	175	5.4	41	9	US-09-984-245-336	Sequence 336, App
29	175	5.4	41	9	US-09-966-262-336	Sequence 336, App
30	175	5.4	41	9	US-09-983-966-336	Sequence 336, App
31	175	5.4	41	9	US-10-143-090-336	Sequence 336, App
32	172	5.3	39	9	US-09-984-245-338	Sequence 338, App
33	172	5.3	39	9	US-09-966-262-338	Sequence 338, App
34	172	5.3	39	9	US-09-983-966-338	Sequence 338, App
35	172	5.3	39	9	US-10-143-090-338	Sequence 338, App
36	166	5.1	38	9	US-09-984-245-333	Sequence 333, App
37	166	5.1	38	9	US-09-966-262-333	Sequence 333, App
38	166	5.1	38	9	US-09-983-966-333	Sequence 333, App
39	166	5.1	38	9	US-10-143-090-333	Sequence 333, App
40	166	5.1	39	9	US-09-984-245-334	Sequence 334, App
41	166	5.1	39	9	US-09-966-262-334	Sequence 334, App
42	166	5.1	39	9	US-09-983-966-334	Sequence 334, App
43	166	5.1	39	9	US-10-143-090-334	Sequence 334, App
44	149	4.6	40	9	US-09-984-245-335	Sequence 335, App
45	149	4.6	40	9	US-09-966-262-335	Sequence 335, App

ALIGNMENTS

RESULT 1
US-09-974-298-184
; Sequence 184, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 184
; LENGTH: 619
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1420940CDI
US-09-974-298-184

Query Match	Score 3216;	DB 9;	Length 619;
Best Local Similarity	98.3%	Pred. No. 2.5e-281;	
Matches 610;	Conservative 2;	Mismatches 2;	Indels 6; Gaps 2;
Oy	1	MSGIKKQKTEKQKSTNVVYQAHVSRNRKQGVGTGRGCTVWLTLGLSGAGKTTISF	60
Db	1	MSGIKKQKTEKQKSTNVVYQAHVSRNRKQGVGTGRGCTVWLTLGLSGAGKTTISF	60
Oy	61	ALAEYLVSHAIPCYSLDGDNVVRHGLNRNLGSPGDREENIRRIAEVAKLFADAGLVCTIS	120
Db	61	ALAEYLVSHAIPCYSLDGDNVVRHGLNRNLGSPGDREENIRRIAEVAKLFADAGLVCTIS	120
Oy	121	FISFPAKDRENAKIHESAGLPFFEIFVDAPLNTICESRDVKGLYKRAGEIKGTGTIDS	180
Db	121	FISFPAKDRENAKIHESAGLPFFEIFVDAPLNTICESRDVKGLYKRAGEIKGTGTIDS	180
Oy	181	DYKPEPTEPVRLKTNLSTVSDCVHQVVELLOPONTIPIVTI IKDHELFPVKNKLDHVAE	240
Db	181	DYKPEPTEPVRLKTNLSTVSDCVHQVVELLOPONTIPIVTI IKDHELFPVKNKLDHVAE	240

Qy		241	AETPLSLSTIKDLQWLQVQLSEGWATPLKGFMRKEYLQVMHFDTLDD-----DGVINMS	295
Db		241	AETPLSLSTIKDLQWLQVQLSEGWATPLKGFMRKEYLQVMHFDTLDDGMLDPDGVINMS	300
Qy		296	IPIVLPVSAEDKTRLEGCSKFVLAHGGRRAILRDAEFYEHRRKEECSRVRWGTTCTCKHPH	355
Db		301	IPIVLPVSAEDKTRLEGCSKFVLAHGGRRAILRDAEFYEHRRKEECSRVRWGTTCTCKHPH	360
Qy		356	IKWMVESGDWLVGGLDQLVLEKIRWNDGLDYRLTPLELKOKCKEMMNADAVFAEQLRNPVH	415
Db		361	IKWMVESGDWLVGGLDQLVLEKIRWNDGLDYRYLTPLELKOKCKEMMNADAVFAEQLRNPVH	420
Qy		416	NGHALLMQDTCRLLLERGYKHPIVLLLHPLGSWTKDDDPVLDWRMKQHAAVLBEGVLDPKS	475
Db		421	NGHALLMQDTRRLLERGYKHPIVLLLHPLGSWTKDDDPVLDWRMKQHAAVLBEGVLDPKS	480
Qy		476	TIVAIFPPSMLYAGTPEVQWHCRSRMIAGANFYITVGRDPAGMPHPETTKDLYBPTHGGKV	535
Db		481	TVAIFPSPMLYAGTPEVQWHCRSRMIAGANFYITVGRDPAGMPHPETTKDLYBPTHGGKV	540
Qy		536	LSMAIPLTSVEIIPRVAAYNKAKAMDVPDPAHRNEFDFISGTRMRKLAREGENPDGCF	595
Db		541	LSMAIPLTSVEIIPRVAAYNKAKAMDVPDPAHRNEFDFISGTRMRKLAREGENPDGCF	600
Qy		596	MAPKAWKVLTYYRSEMDKN	615
Db		601	MAPKAWKVLTYYRS-LEKN	619

RESULT 2

[illegible]

Db 609 KAWTVLTYEYKS 620

RESULT 3

```

; Sequence 196, Application US/09966262
; Publication No. US20030050461A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/966,262
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 196
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-262-196

```

```

Query Match 79.5%; Score 2602; DB 9; Length 624;
Best Local Similarity 77.1%; Pred. No. 6.1e-226;
Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

Qy 5 KKQKTEN-----QOKSTNVVYQAHVSRNKRQGVVGTGGFRGCTVWLTGLSGAGKTTI 58
Db 9 KVKLSNNAQNMGMORATNTVYQAHVSRNKRQGVVGTGGFRGCTVWLTGLSGAGKTTV 68
Qy 59 SFALEEVLVSHAIPCYSLDGNVRHGLNRNLGFSFGDREENIRRIAEYAKIFADAGLVCI 118
Db 69 SMALEEVLVCHGIFCYTLDDGNIRGLNKNLGFSPEDREENVRRIAEYAKIFADAGLVCI 128
Qy 119 TSFTSPFAKORENARKIHESAGLPFFEIFVDAPUNICESRDYKGLYKRRAGEIKGFTGI 178
Db 129 TSFISPTQDRNNARQIHGASLPFFEIFVDAPLHVCEQQRDVKGLYKRRAGEIKGFTGI 188
Qy 179 DSDYEKPEPTELKTNLSTVSDCVHVVVLLQONIVPYTIIDIHLEFVPEKNLDHVR 238
Db 189 DSEYEKPEAPELVLTQSCDVNDVCQVVEILLQERDIVVDASYEVEVKELYVPENKLHLAK 248
Qy 239 AEATLPSLSITKLDLQWVQVLSQWATPLKGFWRKEYLQVHMFDTLDDGVINMSIPI 298
Db 249 TDAETLPALKINKVDMQVQVLAEGWATPLNGFWRREYLOCLHDFCLLDGGINLSYPI 308
Qy 299 VLPVSAEDKTRLEGCSEFVLAHGGRRVAILRDADFEYEHKRECSRVWGTCTTKPHIKM 358
Db 309 VLTATHEDKERLDGCTAFALMYEGRRVAILRNPEFFEHKRECAROWGTCTKNHPYIKM 368
Qy 359 VMESGDMLVGGDLQVLEKIRWNGDLOVRLTPLELKQCKEMNADAVAFOLRNPVHNGH 418
Db 369 VMEQDMLIGDQLQVLDVYVWNGDQLQYRLTPELKQCKEMNADAVAFOLRNPVHNGH 428
Qy 419 ALLMQDTCRRLLERGYKHPVLLHPLGCGTWDVPLDWRMKQHAAYLEEGVLDPKSTIV 478
Db 429 ALLMQDTHKQLLERYGRRPVLLHPLGCGTWDVPLDWRMKQHAAYLEEGVLDPEVTV 488
Qy 479 AIFPSMLYAGPTEVQWCHRSRMATAGANFYIVGRDPAGMPHPETGKDIYEPHSGKVL 538
Db 489 AIFPSMMYAGPTEVQWCHRSRMATAGANFYIVGRDPAGMPHPETGKDIYEPHSGKVL 548
Qy 539 APGLTSVEIIPFRVAAYNAKAKAMDYDPAHNEFDLSGTRMRKLAREGENPPDGFMAP 598
Db 549 APGLTLEIVFRVAAYNAKAKAMDYDSEHDEFEFISGTRMRKLAREGQKPPGCFMAP 608
Qy 599 KAWKVLTDYYS 610
Db 609 KAWTVLTYEYKS 620

RESULT 4
US-09-983-966-196
; Sequence 196, Application US/09983966
; Publication No. US20030060619A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/983,966
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30

```

;; PRIOR APPLICATION NUMBER: US 60/048,350
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,188
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,135
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/050,937
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,187
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,099
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,352
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,186
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,069
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,095
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,131
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,096
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,355
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,160
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,351
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/048,154
;; PRIOR FILING DATE: 1997-05-30
;; PRIOR APPLICATION NUMBER: US 60/054,804
;; PRIOR FILING DATE: 1997-08-05
;; PRIOR APPLICATION NUMBER: US 60/056,370
;; PRIOR FILING DATE: 1997-08-19
;; PRIOR APPLICATION NUMBER: US 60/060,862
;; PRIOR FILING DATE: 1997-10-02
;; NUMBER OF SEQ ID NOS: 343
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 196
;; LENGTH: 624
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-983-966-196

Query Match 79.5%; Score 2602; DB 9; Length 624;
Best Local Similarity 77.1%; Pred. No. 6.1e-226;
Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;
QY 5 KKOKTEN-----QOKSTNVVYQAAHVSRNKRQGVVGTGRGFRGCTVWLTGLSGAGKTTI 58
DB 9 KVKLSNNAQNMGMORATNVYQAAHVSRNKRQGVVGTGRGFRGCTVWLTGLSGAGKTTV 68
QY 59 SFALBEYLVSHPICVSLDGDVNRHGLNRLGFSFGDREENIRRIAEVAKLFADAGLVCI 118
DB 69 SNALEBYLVCHGIPCVTLDDGNIROGLNKLGFSPEDREENVRIAEVAKLFADAGLVCI 128
QY 119 TSFISPFADRENARKIHESAGLPFEIFVDAPLNICESRDVKGLYKARAGEIKGFTGI 178
DB 129 TSFISPYTDORNARQIHESAGLSLPFEFVFDAPLVCEQDVQVYVLLQERDIPVDASYEYKELYVPENKHLAK 188
QY 179 DSDYEKPEPERRVLTNLTSTVSDCVHQVVELLQEQNVPIYTIKDIHELFPVENKLDHVR 238
DB 189 DSEYEKPEAPELVLTDSQVNDVQVVELLQERDIPVDASYEYKELYVPENKHLAK 248
QY 239 AEAETLPSLSITKLDLQVQLSEGWATPLKGFMRKEYLQVNMFTLLDDGVINNSIPI 298
DB 189 DSEYEKPEAPELVLTDSQVNDVQVVELLQERDIPVDASYEYKELYVPENKHLAK 248
QY 239 AEAETLPSLSITKLDLQVQLSEGWATPLKGFMRKEYLQVNMFTLLDDGVINNSIPI 298
DB 249 TDAETLPALKINKVDQMOWVQLAEGWATPLNGFMREYRLQCLHFDCLDDGGVNLSPVI 308
QY 299 VLPVSAEDKTRLEGCSKFLAHGGRVAILRDAEFYEHKREKSRVWGTTCTCKPHIKM 358
DB 309 VLTATHEDEKRLDGGCTAFALMYEGRVAILRNPFEFHKREKRCARQWGTTCCKNHPYIKM 418
DB 358 VMEQDGLWLGGLDQVLDVRYWNGDQYRLTPELQKQKCKEMNADAVFAFQLRNPVHNGH 428

QY 359 VMESGDLVGGDLQVLEKIRWINDGLDQYRLTPELQKQKCKEMNADAVFAFQLRNPVHNGH 418
DB 369 VMEQDGLWLGGLDQVLDVRYWNGDQYRLTPELQKQKCKEMNADAVFAFQLRNPVHNGH 428
QY 419 ALLMQDTCRRLLERGERYKHPVLLHPLGGWTKDDVPLDMRMKQHAHAALVEGVLDPKSTIV 478
DB 429 ALLMQDTHKOLLERGERYRPPVLLHPLGGWTKDDVPLDMRMKQHAHAALVEGVLPETTVV 488
QY 479 AIFPSPMLYAGTEVOMHCHSRMIAGANEYIVGRDPAGMPHPETKKDLYEPHGGVLSM 538
DB 489 AIFPSPMYAGTEVOMHCHSRMVRVAGANFYIVGRDPAGMPHPETGKDLVEPSHGAKVLTM 548
QY 539 APGLTSVEIIPRVAAYNKAAMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAP 598
DB 549 APGLITLIVPFRVAAYNKKKRDYDSEHHEDEFEISGTRMRKLAREGCKPPEGFMAP 608
QY 599 KAWKVLTDIYRS 610
DB 609 KAWTVLTYEYKS 620

RESULT 5
US-10-143-090-196
; Sequence 196, Application US/10143090
; Publication No. US20030069406A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004PI
; CURRENT APPLICATION NUMBER: US/10/143,090
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 196
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-143-090-196

Query Match 79.5%; Score 2602; DB 9; Length 624;
Best Local Similarity 77.1%; Pred. No. 6.1e-226;
Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

QY 5 KKOKTEN-----QOKSTNVVYQAAHVSRNKRQGVVGTGRGFRGCTVWLTGLSGAGKTTI 58
DB 9 KVKLSNNAQNMGMORATNVYQAAHVSRNKRQGVVGTGRGFRGCTVWLTGLSGAGKTTV 68
QY 59 SFALBEYLVSHPICVSLDGDVNRHGLNRLGFSFGDREENIRRIAEVAKLFADAGLVCI 118
DB 69 SNALEBYLVCHGIPCVTLDDGNIROGLNKLGFSPEDREENVRIAEVAKLFADAGLVCI 128
QY 119 TSFISPFADRENARKIHESAGLPFEIFVDAPLNICESRDVKGLYKARAGEIKGFTGI 178
DB 129 TSFISPYTDORNARQIHESAGLSLPFEFVFDAPLVCEQDVQVYVLLQERDIPVDASYEYKELYVPENKHLAK 188
QY 179 DSDYEKPEPERRVLTNLTSTVSDCVHQVVELLQEQNVPIYTIKDIHELFPVENKLDHVR 238
DB 189 DSEYEKPEAPELVLTDSQVNDVQVVELLQERDIPVDASYEYKELYVPENKHLAK 248
QY 239 AEAETLPSLSITKLDLQVQLSEGWATPLKGFMRKEYLQVNMFTLLDDGVINNSIPI 298
DB 249 TDAETLPALKINKVDQMOWVQLAEGWATPLNGFMREYRLQCLHFDCLDDGGVNLSPVI 308
QY 299 VLPVSAEDKTRLEGCSKFLAHGGRVAILRDAEFYEHKREKSRVWGTTCTCKPHIKM 358
DB 309 VLTATHEDEKRLDGGCTAFALMYEGRVAILRNPFEFHKREKRCARQWGTTCCKNHPYIKM 418
DB 358 VMEQDGLWLGGLDQVLDVRYWNGDQYRLTPELQKQKCKEMNADAVFAFQLRNPVHNGH 428

QY 419 ALLMODTCRRLLERYKHPVLLHPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTIV 478
DB 429 ALLMODTHKQLLERYRRPVLHPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTIV 488
QY 479 AIFPSPMLYAGTEVQWCHRCRSMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLMS 538
DB 489 AIFPSPMYAGTEVQWCHRCRSMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLMS 548
QY 539 APLGSVLEIPRVAAYNKKAKMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAP 598
DB 549 APLGLTLEIPRVAAYNKKAKMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAP 608
QY 599 KAWKVLTDYRS 610
DB 609 KAWTVLTYEYS 620

RESULT 6
US-09-771-161A-244
; Sequence 244, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 244
; LENGTH: 624
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-244

Query Match 79.1%; Score 2588; DB 10; Length 624;
Best Local Similarity 76.6%; Pred. No. 1.1e-224;
Matches 469; Conservative 68; Mismatches 69; Indels 6; Gaps 1;

QY 5 KQKOTEN-----QOKSTNVYQAHVSRNKRQGVVGRGGRGCTVWLTGLSGAGKTTI 58
DB 9 KVKLSNNAQNWGMQRATNVTYQAHVSRNKRQGVVGRGGRGCTVWLTGLSGAGKTTI 68
QY 59 SPALAEYLVSHPICYSLDGDNVNRHGLNRNLFSGDREENTRTAEVAKLFADAGLVCI 118
DB 69 SMALEYLVCHPICYTLDGNIROGLNKNLGFSPEDREENVRRRTAEVAKLFADAGLVCI 128
QY 119 TSFISPEAKDRENARKIHESAGLPFFELFVDAPLNICESRDYKGLYKRAAGEIKGFTCI 178
DB 129 TSFISPYQDRNNAKIHESAGLPFFELFVDAPLNICESRDYKGLYKRAAGEIKGFTCI 188
QY 179 DSDYKPEPTEPVLKTNLSTSDCVQVQVVELLQERDIPVDASVEYKELYVPENKHLAK 248
DB 189 DSEYKPEAPELVKLTSDCVNDVCVQVVELLQERDIPVDASVEYKELYVPENKHLAK 248
QY 239 AEAETPLSLITKLDQVQLVSEGWATPLKGFMEKEYLQVHMTDTLLDGVVNMSTPI 298
DB 249 TDAETLPALKINKVDMQVQVLAEGWATPLNGFMREYQLQCLDFCLLDGGVILNSVPI 308
QY 299 VLPVSAEDKTRLEGCSKFLVAGHGRVAILRDABEYHKEKRCRSRVWGTCTTKPHIKM 358
DB 309 VUTATHEKRLDGGTAFALMTEGRRVAILRNPEFFHKEKRCRCARQWGTCTTKNHPYIKM 368
QY 359 VMEQDGLVGGDLQVLEIRWNGDGLDQYRLTPELQKCKEMNADAVAFQLRNPNVHNGH 418
DB 369 VMEQDGLVGGDLQVLEIRWNGDGLDQYRLTPELQKCKEMNADAVAFQLRNPNVHNGH 428

QY 419 ALLMODTCRRLLERYKHPVLLHPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTIV 478
DB 429 ALLMODTHKQLLERYRRPVLHPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTIV 488
QY 479 AIFPSPMLYAGTEVQWCHRCRSMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLMS 538
DB 489 AIFPSPMYAGTEVQWCHRCRSMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLMS 548
QY 539 APLGSVLEIPRVAAYNKKAKMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAP 598
DB 549 APLGLTLEIPRVAAYNKKAKMDFYDPAHNEFDFISGTRMRKLAREGENPPDGFMAP 608
QY 599 KAWKVLTDYRS 610
DB 609 KAWTVLTYEYS 620

RESULT 7
US-10-156-239-40
; Sequence 40, Application US/10156239
; Publication No. US20030036074A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: No. US20030036074A1el Nucleic Acid Sequences Encoding Human Tr
; TITLE OF INVENTION: Appase Molecule, A Human Ubiquitin Hydrolase-Like Molecule, A
; TITLE OF INVENTION: Ubiquitin Conjugating Enzyme-Like Molecule, and Uses Therefor
; FILE REFERENCE: 35800/247645
; CURRENT APPLICATION NUMBER: US/10/156,239
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 09/7795,693
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/185,906
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 09/809,557
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/192,018
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,568
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/191,790
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 09/808,767
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/191,781
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pfam consensus sequence
US-10-156-239-40

Query Match 17.5%; Score 571.5; DB 9; Length 162;
Best Local Similarity 67.9%; Pred. No. 1.1e-43;
Matches 110; Conservative 17; Mismatches 32; Indels 3; Gaps 1;

QY 41 RCTVWLTGLSAGKTTISFALAEYLVSHPICYSLDGDNVNRHGLNRNLFSGDREENT 100
DB 1 RCTVWLTGLSAGKTTISFALAEYLVSHPICYSLDGDNVNRHGLNRNLFSGDREENT 100
QY 101 RRIAEVAKLFADAGLVCIISFISPAKORENARKTH---ESAGLPFFELFVDAPLNICES 157
DB 61 RRVGEVAKLFADAGLVCIISFISPAKORENARKTH---ESAGLPFFELFVDAPLNICES 157
QY 158 RRVKGLYKRAAGEIKGFTGIDSDYKPEPTEPVLKTNLSTV 199
DB 121 RDPKGLYKRAAGEIKGFTGIDSDYKPEPTEPVLKTNLSTV 199


```

, , PRIOR APPLICATION NUMBER: US 09/632,366
, , PRIOR FILING DATE: 2000-08-03
, , PRIOR APPLICATION NUMBER: GB 24263.6
, , PRIOR FILING DATE: 2000-10-04
, , PRIOR APPLICATION NUMBER: US 60/236,359
, , PRIOR FILING DATE: 2000-09-27
, , PRIOR APPLICATION NUMBER: PCT/US01/006666
, , PRIOR FILING DATE: 2001-01-30
, , PRIOR APPLICATION NUMBER: PCT/US01/006667
, , PRIOR FILING DATE: 2001-01-30
, , PRIOR APPLICATION NUMBER: PCT/US01/006668
, , PRIOR FILING DATE: 2001-01-30
, , PRIOR APPLICATION NUMBER: PCT/US01/006669
, , PRIOR FILING DATE: 2001-01-30
, , PRIOR APPLICATION NUMBER: PCT/US01/006670
, , PRIOR FILING DATE: 2001-01-30
, , PRIOR APPLICATION NUMBER: PCT/US01/006671
, , PRIOR FILING DATE: 2000-09-21
, , PRIOR APPLICATION NUMBER: US 09/608,408
, , PRIOR FILING DATE: 2000-06-30
, , PRIOR APPLICATION NUMBER: US 09/774,203
, , PRIOR FILING DATE: 2001-01-29
, , NUMBER OF SEQ ID NOS: 49117
, , SOFTWARE: Annomax Sequence Listing Engine
, , SEQ ID NO 39981
, , LENGTH: 65
, , TYPE: PRMT
, , ORGANISM: Homo sapiens
, , FEATURE:
, , OTHER INFORMATION: MAP TO AC006191.1
, , OTHER INFORMATION: EXPRESSED IN FETAL I
, , OTHER INFORMATION: EXPRESSED IN BONE MAR
, , OTHER INFORMATION: EXPRESSED IN PLACENT
, , OTHER INFORMATION: EXPRESSED IN LUNG, S
, , OTHER INFORMATION: EXPRESSED IN HEART,
, , OTHER INFORMATION: EXPRESSED IN BRAIN,
, , OTHER INFORMATION: EXPRESSED IN ADULT I
, , OTHER INFORMATION: SWISSPROT HIT: 09534
, , OTHER INFORMATION: EST_HUMAN HIT: BE299
, , US-09-864-761-39981

```

```

; CURRENT FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 635
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-391-635

Query Match          9.9%; Score 323; DB 9; Length 174;
Best Local Similarity 83.3%; Pred. No. 3.le-21;
Matches 65; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY   1 MSGIKKQKTENQOKSTNVVYQAHHVSNNKRGVVGTGGFRGCTVWLTGLSGAGKTTISF 60
      |||||
Db    55 MSGIKKQKTENQOKSTNVVYQAHHVSNKRSGVVGTGGFRGCTVWLTGLSGAGKNNDKF 114
      |||||

QY   61 ALEEVLVSHAIPCYSLDG 78
      |||||
Db   115 CPGGVLVSHAIPVNSWMG 132
      |||||

RESULT 13
US-09-867-550-96
; Sequence 96, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 96
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-96

Query Match          6.8%; Score 221.5; DB 10; Length 163;
Best Local Similarity 41.9%; Pred. No. 4.le-12;
Matches 57; Conservative 17; Mismatches 59; Indels 3; Gaps 2;

QY   41 RCGTVWLTGLSGAGKTTISFALEEVLSHAIPCYS-LDGNVRHGLNRNLGFSPGDREEN 99
      :||:||||:||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db   26 QGFTLLSGLNSGKDQIASALQVKHQGRSVSLLDGTVRHEJSSSELGFGSPEDRRHN 85
      :|||||
QY   100 IRIAEVAKLFADAGLVCIIFSIFPAKDRENARK--IHESAGLPFFEIFVDAPLNICES 157
      :|||||
Db    86 LQRIAFVASBELTRAGAIVASPIAPIEHSQMARDTILKGGSNNFLIHVATPFEICEK 145
      :|||||

QY   158 RDVKGLYKRRAGEIK 173
      |::|::|::|
Db   146 HDRRGNYAKARRGEVK 161
      |::|::|::|

RESULT 14
US-09-984-245-340
; Sequence 340, Application US/09984245
; Patent No. US20020165374A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1

```


Db :|||||||||||||:||||||||||||||||||||||||||||| |||
1 MYAGTEVQWHCRAHWAGANFYIVGRDPAGMHPETGKDL 41

Search completed: May 30, 2003, 12:28:53
Job time : 49 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:15:39 ; Search time 43 Seconds
(without alignments)
1374.946 Million cell updates/sec

Title: US-09-898-165B-7

Perfect score: 3271

Sequence: 1 MSGIKKQKTKNQKSTNVVY.....MAPKAMKVLTDYRSEMDKN 615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2602	79.5	624	1 JW0087	adenyl-yl-sulfate k
2	2367	72.4	610	1 JC4383	adenyl-yl-sulfate k
3	1887	57.7	652	1 T24918	3'-phosphoadenosin
4	1299	39.7	424	2 S44079	sulfate adenyl-yltr
5	1282	39.2	465	2 E71409	sulfate adenyl-yltr
6	1274	38.9	463	2 S44267	sulfate adenyl-yltr
7	1268.5	38.8	476	2 S44943	sulfate adenyl-yltr
8	1265	38.7	483	2 T14475	sulfate adenyl-yltr
9	1259.5	38.5	469	2 T52659	sulfate adenyl-yltr
10	1230.5	37.6	489	1 T01204	sulfate adenyl-yltr
11	1212.5	37.1	459	2 T08594	probable sulfate a
12	556	17.0	202	2 T50101	adenyl-yl-sulfate k
13	551.5	16.9	200	2 E96912	adenyl-yl-sulfate k
14	523	16.0	202	1 S17244	adenyl-yl-sulfate k
15	500	15.3	199	2 A69839	adenyl-yl-sulfate k
16	497.5	15.2	293	2 T06100	adenyl-yl-sulfate k
17	493.5	15.1	213	2 A70408	adenyl-yl-sulfate k
18	487.5	14.9	202	2 A83836	adenyl-yl-sulfate k
19	481.5	14.7	202	2 A10856	adenosine 5'-phosp
20	481.5	14.7	635	2 A87433	hypothetical prote
21	476.5	14.6	201	1 B65056	adenyl-yl-sulfate k
22	476.5	14.6	201	2 D51079	adenosine 5'-phosp
23	476.5	14.6	201	2 E85924	adenosine 5'-phosp
24	476	14.6	276	1 S47640	adenyl-yl-sulfate k
25	469.5	14.4	312	2 T08076	adenyl-yl-sulfate k
26	463	14.2	215	2 F82062	adenyl-yl-sulfate k
27	456.5	14.0	614	1 B70772	probable adenyl-yl
28	451.5	13.8	197	2 C59877	adenyl-yl-sulfate k
29	440	13.5	644	2 AD3471	adenyl-yl-sulfate k

ALIGNMENTS

RESULT 1

JW0087

adenyl-yl-sulfate kinase (EC 2.7.1.25) - human

N:Alternate names: adenosine 5'-phosphosulfate kinase; PAPS

N:Contains: adenylsulfate kinase (EC 2.7.1.25); sulfate adenyltransferase (EC 2.7

C:Species: Homo sapiens (man)

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 03-Jun-2002

C:Accession: JW0087

R:Yanagisawa, K.; Sakakibara, Y.; Suiko, M.; Takami, Y.; Nakayama, T.; Nakajima, H.

Biosci. Biotechnol. Biochem. 62, 1037-1040, 1998

A:Title: cDNA cloning, expression, and characterization of the human bifunctional ATP

A:Reference number: JW0087; MUID:98312048; PMID:9648242

A:Accession: JW0087

A:Molecule type: mRNA

A:Residues: 1-624 <YAN>

A:Cross-references: GB:AR033026; NID:93378100; PIDN:AAC28429.1; PID:93378101

A:Experimental source: Brain

C:Function: <ASKF>

C:Function: as adenylsulfate kinase catalyzes the phosphorylation of adenylisul

C:Function: <SATE>

C:Description: as sulfate adenyltransferase catalyzes the reaction of sulfate and A

C:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenylsulfat

C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loop

F:52-215/Domain: adenylsulfate kinase homology <ASK>

F:59-66/Region: nucleotide-binding motif A (P-loop)

F:226-620/Domain: sulfate adenyltransferase homology <SAT>

F:133/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 79.5% Score 2602; DB 1; Length 624;

Best Local Similarity 77.1%; Pred. No. 6.4e-186;

Matches 472; Conservative 67; Mismatches 67; Indels 6; Gaps 1;

QY 5 KKQKTE-----QKSTNVVYQAHVSRNKRKGGVGTGGRGCTVWLTGLSGAGKTTI 58

DB 9 KKLUSNAQNMGMQRATNVITYQAHVSRNKRKGGVGTGGRGCTVWLTGLSGAGKTTV 68

QY 59 SFALFEYLVSHAIPCYSLDGDNRVHGLNRNLGFSFGDREENTRIRAEVAKLFADAGLVCI 118

DB 69 SMALEEYLVCHGICPYCTILDGNIROGLNKNLGNLGSFDPEDREENVRRIRAEVAKLFADAGLVCI 128

QY 119 TSFTSPFAKDRENARKIHESAGLPFEIFVDAPLNCESROVKGLYKARAGEIKGFTGI 178

DB 129 TSFISPYTQDRNARQIHEGASLPFEFVFDAPLVCEQDQVKGGLYKARAGEIKGFTGI 188

QY 179 DSDVEKPTPEVLTNLSTVSDCVHVVQLQEQNIPTTIKDIHELFPENKLDHVR 238

DB 189 DSEYKPEAPELVLTDLQVQLSEGWATPLKGFMEKEYLQVHFDFTLLDDGVINSLKLA 248

QY 239 AEATFLPSLSTTKLDLQVQLSEGWATPLKGFMEKEYLQVHFDFTLLDDGVINSLKLA 298

DB 249 TDAETPLPAKINKVDNQVQLSEGWATPLNGFMREYVQLQCLHFDCLLDGVINSLVPI 308

probable adenyl-yl-
adenyl-yl-sulfate k
adenyl-yl-sulfate k
adenyl-yl-sulfate k
adenyl-yl-sulfate k
adenyl-yl-sulfate k
ATP sulfurylase, l
probable adenyl-yl-
adenyl-yl-sulfate k
adenosine 5'-phosp
sulfate adenyl-yltr
adenyl-yl-sulfate k
sulfate adenyl-yltr
sulfate adenyl-yltr
sulfate adenyl-yltr
sulfate adenyl-yltr

QY 299 VLPVSAEDKTRLEGGCKSFVLAHGGRRVAILRDAAEFYEHKREKCSRVMGTTCTKHPHIKM 358
Db 309 VLTATHEDKRLDGCATAFALMYEGRRVAILRNPEFFEHKREKCAROMGTCTCKNHPYIKM 368
QY 359 VMESGDMVLGGDLQVLEKIRWNGDLDQYRLTPELKOKCKEMNADAVAFQLRNPVHNGH 418
Db 369 VMESGDMVLGGDLQVLDRTVWNGDLDQYRTPTELKQKFKDMNADAVAFQLRNPVHNGH 428
QY 419 ALLMODTCRRLLRGYKHPVLLHPLGGWTKDDVPLDWRMKQAAVLEEGVLDPKSTIV 478
Db 429 ALLMODTHKQLLGRYRRPVLHPLGGWTKDDVPLDWRMKQAAVLEEGVLDPKSTIV 488
QY 479 AIFPSPMLYAGPTEVQVHCRSMIAGANFYIVGRDPAGMPHETKDYEPHGGKVLMS 538
Db 489 AIFPSPMLYAGPTEVQVHCRSMIAGANFYIVGRDPAGMPHETKDYEPHGGKVLMS 548
QY 539 APLGTSVEIIPFVAAYNAKAMDFYDPAHNEFDFISGTRMKRLAREGENPDDGFMAP 598
Db 549 APLGTSVEIIPFVAAYNAKAMDFYDPAHNEFDFISGTRMKRLAREGENPDDGFMAP 608
QY 599 KAWKVLTDYYS 610
Db 609 KAWKVLTDYYS 620

RESULT 2
JC4383
adenyl-yl-sulfate kinase (EC 2.7.1.25) - spoonworm (Urechis caupo)
N:Alternate names: adenosine 5'-phosphosulfate kinase; PAPS
N:Contains: adenyl-yl-sulfate kinase (EC 2.7.1.25); sulfate adenyl-yl-transferase (EC 2.7.7.4)
C:Species: Urechis caupo
C:Date: 17-Jan-1996 #sequence_revision 19-Apr-1996 #text_change 03-Jun-2002
C:Accession: JC4383
R:Rosenthal, E.; Leustek, T.
Gene 165, 243-248, 1995
A:Title: A multifunctional Urechis caupo protein, PAPS synthetase, has both ATP sulfurylase and adenylyl-sulfate kinase activities
A:Reference number: JC4383; MUID:96096529; PMID:8522184
A:Accession: JC4383
A:Molecule type: mRNA
A:Residues: 1-610 <ROS>
A:Cross-references: GB:I39001; NID:g705384; PIDN:AAB00139.1; PID:g705385
C:Genetics:
A:Gene: paps
A:Function: <ASKF>
A:Description: as sulfate adenyl-yl-sulfate kinase catalyzes the phosphorylation of adenyl-yl-sulfate
C:Function: <SAIF>
A:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenyl-yl-sulfate
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loop;
F:37-200/Domain: adenyl-yl-sulfate kinase homology <ASK>
F:44-51/Region: nucleotide-binding motif A (P-loop)
F:161-171/Domain: 3'-phosphoadenosine-5'-phosphosulfate binding #status predicted <PAB>
F:211-605/Domain: sulfate adenyl-yl-transferase homology <SAT>
F:118/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 72.4%; Score 2367; DB 1; Length 610;
Best Local Similarity 71.4%; Pred. No. 2.1e-168;
Matches 432; Conservative 69; Mismatches 104; Indels 0; Gaps 0;

QY 11 NOCKSTNVVYQAHVSRNKGQVGRGGFRGCTVMTGLSGAGKTTISFALFEYLVS 70
Db 6 NGOLATNVTFQTHVSRNKGQVGRGGFRGCTVMTGLSGAGKTTISFALFEYLVS 65

QY 71 IPCYSLDGNVRHGLNRLNFGSPGDRNIRIAEVAKLFDAGLVLCITISFPAKDR 130
Db 66 IPTYSLDGNVRHGLNRLNFGSPGDRNIRIAEVAKLFDAGLVLCITISFPAKDR 125

QY 131 NARKIHESAGLPEFFEDVAPLNCESRDVKGILKARAGEIKGFGIDSDYKPTPR 190
Db 126 LARSLHEQAGLPEFFEDVAPLNCESRDVKGILKARAGEIKGFGIDSDYKPTPR 185

QY 191 VLKTNLSJSTDCVQHVVELLQEQNIYPTTIKIDIELFYPENKLDHVAEAEFLPSLIT 250

Db 186 QLYVAGNKSDICVQEVVVVLLQKNGVVPESAVNIVKELFVPESGLEHAKAEIVDLPMTMEIT 245
QY 251 KLDLOWQVLSSEGWATPLKGFMEKEKYLQVMHFDFTLLDDGVINMSIPVLPVSAEDKTRL 310
Db 246 KLDLOWQVLSSEGWATPLKGFMEKEKYLQVMHFDFTLLDDGVINMSIPVLPVSAEDKTRL 305
QY 311 EGCSKFLVLAHGGRRVAILRDAAEFYEHKREKCSRVMGTTCTKHPHIKMVMSGDMVLGG 370
Db 306 EGCSKFLVLAHGGRRVAILRDAAEFYEHKREKCSRVMGTTCTKHPHIKMVMSGDMVLGG 365
QY 371 LQVLEKIRWNGDLDQYRLTPELKOKCKEMNADAVAFQLRNPVHNGHALLMODTCRRLL 430
Db 366 LEVLEHRTWNGDLDQYRLTPELKOKCKEMNADAVAFQLRNPVHNGHALLMODTCRRLL 425
QY 431 ERGYKHPVLLHPLGGWTKDDVPLDWRMKQAAVLEEGVLDPKSTIVAFSPMLYAGP 490
Db 426 ERGYKHPVLLHPLGGWTKDDVPLDWRMKQAAVLEEGVLDPKSTIVAFSPMLYAGP 485
QY 491 TEVQVHCRSMIAGANFYIVGRDPAGMPHETKDYEPHGGKVLMSAPGLTSVEIIPF 550
Db 486 TEVQVHCRSMIAGANFYIVGRDPAGMPHETKDYEPHGGKVLMSAPGLTSVEIIPF 545
QY 551 RVAAYNKKAKAMDFYDPAHNEFDFISGTRMKRLAREGENPDDGFMAPKAWKVLTDYYS 610
Db 546 RVAAYNKKAKAMDFYDPAHNEFDFISGTRMKRLAREGENPDDGFMAPKAWKVLTDYYS 605
QY 611 EMDKN 615
Db 606 KAQOS 610

RESULT 3
T24918
3'-phosphoadenosine-5'-phosphosulfate synthetase - Caenorhabditis elegans
N:Alternate names: protein T14G10.1
C:Species: Caenorhabditis elegans
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 19-Jan-2001
C:Accession: T24918
R:Wild, A.
Submitted to the EMBL Data Library, January 1996
A:Reference number: Z19954
A:Accession: T24918
A:Status: preliminary; translated from CB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-652 <WIL>
A:Cross-references: EMBL:Z68880; PIDN:CAA93098.1; GSPDB:GN00022; CESP:T14G10.1
A:Experimental source: clone T14G10
C:Genetics:
A:Gene: CESP:T14G10.1
A:Map position: 4
A:Introns: 23/3; 82/3; 176/1; 535/2; 623/2
C:Function: <ASKF>
A:Description: as adenyl-yl-sulfate kinase catalyzes the phosphorylation of adenyl-yl-sulfate
C:Function: <SAIF>
A:Superfamily: animal 3'-phosphoadenosine-5'-phosphosulfate synthetase; adenyl-yl-sulfate
C:Keywords: multifunctional enzyme; nucleotide binding; nucleotidyltransferase; P-loop;
F:36-219/Domain: adenyl-yl-sulfate kinase homology <ASK>
F:56-70/Region: nucleotide-binding motif A (P-loop)
F:180-190/Domain: 3'-phosphoadenosine-5'-phosphosulfate binding #status predicted <PA>
F:230-647/Domain: sulfate adenyl-yl-transferase homology <SAT>
F:137/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 57.7%; Score 1887; DB 1; Length 652;
Best Local Similarity 56.5%; Pred. No. 1.4e-132;
Matches 362; Conservative 100; Mismatches 151; Indels 28; Gaps 6;

QY 1 MSGIKKQKTEEN-QOKSTNVVYQAHVSRNKGQVGRGGFRGCTVMTGLSGAGKTTIS 59
Db 14 MPMLKPRVSSLSGOSTNITYEHTLSREERAAVGRHEGFRGCTVMTGLSGAGKTTIS 73

QY 60 FALFEYLVSHTAIPCYSLDGNVRHGLNRLNFGSPGDRNIRIAEVAKLFDAGLVLCIT 119

Db 74 FALERTLANKIGPCYGLDGNIRHGLCKNLGFSKEDRQBNIRRAEVAKLFADSGMICIA 133
QY 120 SFISPAKRENAKTHESAGLPFEFVDAPLNICESRDVGLYKRARAGETKGTGID 179
Db 134 AFISPOEDRLARKTHESENVKFEVHVSTTLEVCEDQRPQOLYKARAGLGTGID 193
QY 180 SYEKPEPTEPVLKNTLSVDCVHQVVELLOEQNIVPTI--IKDIHFLVFPEN-KLDH 236
Db 194 SAYEPENAEIILDAGKQCVQCVKVLHESKGLLPEQIPDPVPAVRELFVSDDLTVAE 253
QY 237 VRAEATLPSLTKLDLQWQVLSGFWATPLKGFMRKEYLOVHMFDTLID----- 288
Db 254 LKESONLPTVELTKVDLQWQVLSGFWATPLKGFMRKEYLOVHMFDTLIDLKHKVAFV 313
QY 289 -----DGVINMSPIVLPVSAEDKTRLEGCSKFVLHGGRRVAILRDAEFY 334
Db 314 GEKSDDKEDSWPMDD-INGSPVLPISDDVKKLEGVTRALKYNGQVYALLSPEIF 372
QY 335 EHRKEBRCRVGTTCTKPHIKMWNESGDWLVGGDLQVLEKIRWNGDLQVRLTPLELK 394
Db 373 EHRKDRVCROFTNDPRHPAQAQVLESGNWLGGDVAVVQKIQFNGDLQVRYKTPNELR 432
QY 395 QCKEKMNADAVAFOLRNPHNGHALLMODTCRLLERGYKHPVLLHPLGGWTKDDVP 454
Db 433 AIPAENADAVAFOLRNPHNGHALLMODTCRLLERGYKHPVLLHPLGGWTKDDVP 491
QY 455 LDWRMKAHAALVEEGVLDPKSTIVAFPPSMYAGTPEVQWCHRSRMIAGANFYIVGRDP 514
Db 492 LDIRIKQHEAVTAERVLDPENWVLSIFPSPMYAGTPEVQWCHRSRMIAGANFYIVGRDP 551
QY 515 AGMHPETKDYEPHGGKVLSPMAPGLTSVEIIPRVAAYNKAAMFYDPAHNEFD 574
Db 552 AGIQKPGSDALYETHGAKVLSMAPGLTSVEIIPRVAAYNKAAMFYDPAHNEFD 611
QY 575 FTSCTMRKLAREGENPPDGFMAPKAWKVLTDYRSEMDKN 615
Db 612 NISGTMKRLARNGDTPPEGFMAPTAEWLAGYKSLQNS 652

RESULT 4
S44079
sulfate adenylyltransferase (EC 2.7.7.4) met3-1 - potato
C:Species: Solanum tuberosum (potato)
C:Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
R:Accession: S67497; S44079
R:Klonus, D.; Hoefgen, R.; Willmitzer, L.; Riesmeyer, J.W.
Plant J. 6, 105-112, 1994
A:Title: Isolation and characterization of two cDNA clones encoding ATP-sulfonylases from
A:Reference number: S67497; MUID:95004649; PMID:7920699
A:Accession: S67497
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-424 <KLO>
A:Cross-references: EMBL:X75041; NID:9471342; PIDN:CAA52953.1; PID:g531495
C:Genetics:
A:Gene: met3-1
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology
C:Keywords: nucleotidyltransferase
F:11-409/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 39.7%; Score 1299; DB 2; Length 424;
Best Local Similarity 61.7%; Pred. No. 5.1e-89;
Matches 240; Conservative 64; Mismatches 81; Indels 4; Gaps 4;

QY 226 ELFPVKNKLDHVRAREATLPSLITKLDLQWQVLSGFWATPLKGFMRKEYLOVHMFD 285
Db 21 ELFVDESQDLKREKALNPKIKLTKIDVWVHVLSGFWATPLKGFMRKEYLOVHMFD 80
QY 286 L-LDGC-VINMSPIVLPVSAEDKTRLEGCSFVL-AHGGRVAILRDAEFYHREK 342
Db 81 IRLDGSVNMSPVILADDDSKQKQIDGSSVALVVDGNNPAILTDIEYKHKEER 140
QY 343 SRWGTCTCTKPHI-KMWNESGDWLVGGDLQVLEKIRWNGDLQVRLTPLELKQCKEMN 401

Db 141 ARTWGTAPCLPVVDEAITHSGNWLKGLVTEPKYKHGDLDSFRLSPSELRAEFTRRN 200
QY 402 ADAVFAFOLRNPHNGHALLMODTCRLLERGYKHPVLLHPLGGWTKDDVDPLDWRMK 461
Db 201 ADAVFAFOLRNPHNGHALLMODTCRLLERGYKHPVLLHPLGGWTKDDVDPLDWRMK 260
QY 462 HAAVLEEGVLDPKSTIVAFPPSMYAGTPEVQWCHRSRMIAGANFYIVGRDPAGMPHE 521
Db 261 HEKVLGEGVLDPEVTLVSIFFSPMYAGTPEVQWCHRSRMIAGANFYIVGRDPAGMPHE 320
QY 522 TKDLYEPHGGKVLSPMAPGLTSVEIIPRVAAYNKAAMFYDPAHNEFDIFSTRM 581
Db 321 EKRDLVDADHGGKVLSPMAPGLTSVEIIPRVAAYNKAAMFYDPAHNEFDIFSTRM 380
QY 582 RKLAREGENPPDGFMAPKAWKVLTDYRSEMDKN 610
Db 381 RTLAUKTESPPDGFMAPKAWKVLTDYRSEMDKN 409

RESULT 5
E71409
sulfate adenylyltransferase (EC 2.7.7.4) precursor (clone APS3) - Arabidopsis thaliana
N:Alternate names: ATP sulfurylase
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
R:Accession: E71409; S68201
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.;
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenger, T.; Pohl, T.M.; Terry, N.;
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk,
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.;
Rechman, S.;
Chalatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: E71409
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-465 <BEV>
A:Cross-references: GB:297336; NID:g2244788; PIDN:CAB10247.1; PID:g2244824
R:Murillo, M.; Leustek, T.
Arch. Biochem. Biophys. 323, 195-204, 1995
A:Title: Adenosine-5'-triphosphate-sulfurylase from Arabidopsis thaliana and Escheri
ine-5'-triphosphate-sulfurylase cDNAs from Arabidopsis thaliana and analysis of a
A:Reference number: S68024; MUID:96019964; PMID:7487067
A:Accession: S68201
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-6,'D',8-35,'A',37-465 <MUR>
A:Cross-references: EMBL:U06275; NID:g459143; PIDN:AAA92350.1; PID:g459144
C:Genetics:
A:Map position: 4COP9-4G3845
A:Gene: nuclear
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homol
C:Keywords: chloroplast; nucleotidyltransferase
F:1-50/Domain: transit peptide (chloroplast) #status predicted <TNP>
F:51-465/Product: sulfate adenylyltransferase #status predicted <MAT>
F:52-450/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 39.2%; Score 1282; DB 2; Length 465;
Best Local Similarity 62.3%; Pred. No. 1.1e-87;
Matches 243; Conservative 60; Mismatches 81; Indels 6; Gaps 5;

QY 226 ELFPVKNKLDHVRAREATLPSLITKLDLQWQVLSGFWATPLKGFMRKEYLOVHMFD 285
Db 62 DLVVKPRRREKHEADLPVRLTAIDLQWVLSGFWATPLKGFMRKEYLOVHMFD 121
QY 286 L-LDGC-VINMSPIVLPVSAEDKTRLEGCSFVL-AHGGRVAILRDAEFYHREK 341
Db 122 LNLDDGSVNMSPVILADDDSKQKQIDGSSVALVVDGNNPAILTDIEYKHKEER 180

```

342 QY CSRVVGTCTCKPHI - KMWESGDMVLVGGDLQVLEKIRWNGDLQVRLTPTLELKQKCKEM 400
      ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : :
181 Db IARTWGTAPGLPYVEEAITNAGDMVLVGGDLVLEPKVYNGDLDRFLRPFELRKELEKR 240
      ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : :
401 QY NADAVFAQLRPNVHNGHALLMODTCRRLLERGYKHPVLLHPLGCGTTKDDDVPLDWRMK 460
      ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : :
241 Db GADAVFAQLRPNVHNGHALLMTDRRLLEMGYKNPILLHPLGGFTKADVPLPSWRMK 300
      ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : :
461 QY QHAAVLEGVLDPKSTTVAIFPSPMLYAGTVEVQWECRSMTAGANFYIVGRDPPAGMHPH 520
      ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : :
301 Db QHEKVEGVLDPKETTVVIFPSPMLYAGTVEVQWHAKEKINAGANFYIVGRDPPAGMHPH 360
      ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : :
521 QY ETCKDLYEPHTGGKVLVSNAPGLTSVEIIPRVAAYNKAKKAMDFYDPARHNEFDTSIGTR 580
      ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : :
361 Db VEKRDLYDADHGKVLVSNAPGLERLNLIPRVAAYDKTQGMKAFDPDSRAQDFLFIGTK 420
      ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : :
581 QY MRKILAREGENPPDGMFAPKAWKVLTDYYRS 610
      ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : :
421 Db MRALAKNRENPPDGMFCMPCGKVKLVDDYDS 450
      ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : : ||||| : : : :

RESULT 6
S44267
sulfate adenylyltransferase (EC 2.7.7.4) met3-2 - potato
C:Species: Solanum tuberosum (potato)
C:Date: 08-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C:Accession: S67498; S44267
R:Klonus, D.; Hoeftgen, R.; Willmitzer, L.; Riesmeier, J.W.

```

RESULT 7

S44943

sulfate adenylyltransferase (EC 2.7.7.4) met3-1 precursor - Arabidopsis thaliana

N:Alternate names: ATP sulfurylase; sulfate adenylyltransferase APS2

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 06-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999

C:Accession: S44943; S68024; S68202

R:Klonus, D.

submitted to the EMBL Data Library, May 1994

A:Description: ATP-sulfurylase cDNA clone from A.thaliana.

A:Reference number: S44943

A:Accession: S44943

A:Molecule type: mRNA

A:Residues: 1-476 <KLO>

A:Cross-references: EMBL:X79210; NID:G487403; PIDN:CAA55799.1; PID:G487404

R:Murillo, M.; Leustek, T.

Arch. Biochem. Biophys. 323, 195-204, 1995

A:Title: Adenosine-5'-triphosphate-sulfurylase from Arabidopsis thaliana and Escherichia coli

A:Reference number: S68024; MUID:96019564; PMID:7487067

A:Accession: S68024

A:Molecule type: mRNA

A:Residues: 1-476 <MUW>

A:Cross-references: EMBL:U06276; NID:G1228103; PIDN:AAA92351.1; PID:G1228104

A:Experimental source: clone APS2

C:Genetics:

A:Gene: met3-1

A:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homolog

C:Keywords: Chloroplast; nucleotidyltransferase

F:1-62/Domain: transit peptide (chloroplast) #status predicted <NP>

F:63-476/Product: sulfate adenylyltransferase met3-1 #status predicted <MAT>

F:64-462/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 38.8%; Score 1268.5; DB 2; Length 476;

Best Local Similarity 57.3%; Pred. No. 1.le-86;

Matches 242; Conservative 71; Mismatches 90; Indels 19; Gaps 6;

QY 207 VELQEQNIVPVTIK-----DIHELFPENKLDHVAEETLPSLSITKLD 253

DB 42 LNLVYKRLNTWQSVSKMTVKSSLDPDGGELVELIVPTEIGVKASETMPKVKLNQID 101

QY 254 LQWQVLSEGWATPLUGKFMREKEYLQVNHFDL-LDDGV-INMSIPIVLVPSAEDKTRLE 311

DB 102 LEWHVISEGWASPLUGFMREDEYQLSHFNLSRLKNGTFVMSLPIVLAIIDDDTKEQI- 160

QY 312 GCSKFEV--LAHGRRVAILROAEFYEHRKEKRCRWGTTCTKHPHI-KMYMESDMLVG 368

DB 161 GGSSENVALVPCOGDITIGLSRSEIYKHNKEARIATGTTSGLPYVEEYITPSSGNWLG 220

QY 369 GDVLQLEKIRWNGDLQXRLTPLELKQKCKEMNADAVEAFQLRNPNVHGHALLMDTCRR 428

DB 221 GDLEVFEPKYNGLDHLVLSFKQLREEDFNQADAVEAFQLRNPNVHGHALLMNDTRK 280

QY 429 LLERGYKHPVLLHPLGGWTKDDVPLDWRMKQHAAYLVEEGVLDPKSTIVAFIPSPMLYA 488

DB 281 LLEMGKKNPVLLHPLGGFTKADDDVPLDVRMEQHSKVLGDDGVLDPKTTIVSIFSPMHYA 340

QY 489 GPTVEQWCHRSRMTAGANFYIVGRDPACGMPIPETKDLIYETHGKGLVSMAPGLTSVEII 548

DB 341 GPTVEQWHAIRINAGANFYIVGRDPACGMGHPTEKRLDYDPDHGKRVLSMAPGLEKLNIL 400

QY 549 PFRVAAYNKAQKAMDYDPARHNEPDFISGTRMRKLAREGENPPDGFVAPKAWKVLTDY 608

DB 401 PFRVAAYDTIEKKMAFFDPSRAKEFLFISGTRMRYARTGENPPDGFVAPKAWKVLTDY 460

QY 609 RS 610

DB 461 ES 462

Query Match 37.6%; Score 1230.5; DB 1; Length 489;
Best Local Similarity 59.1%; Pred. No. 8.1e-84;
Matches 233; Conservative 66; Mismatches 86; Indels 9; Gaps 6;

QY 226 ELFVPE--NKLDHVRRAEATLP--SLSTTKLDLQWQVLSSEGWATPLKGFMRKEYLEQVMH 282
DB 80 DLVAPEGGRRAALRRRAELPHRLRUGRVKWHVLSSEGWASPLOGFMRDDEFFQTLH 139
QY 283 FDTL--LDDGVINMSIPVLVPSAEDKTRL--EGCSKF--VLHGGRRVAILRDAEFYEH 337
DB 140 FNAIRGODGRWNVMSVIVLSVGDQORRAIQADGATRAVAVDEDRPIAVLSIDIEYKHN 199
QY 338 KEERCSTWGTCTTKPHI--KMWESGDWLVGGDLQVLEKIRNWDGLDQYRLTPLELKQK 396
DB 200 KEERIARTWGTATAPLVVEEAITNAGDWLIGDLEVPKYNKNDGLDQYRLSPAQLREE 259
QY 397 CKENADAVFAFOLRNPVHNGHALLMDTCRRLLERGYKHPVLLHPLGGTWDKDDVPLD 456
DB 260 FARNADAVFAFOLRNPVHNGHALLMDTCRRLLERGYKHPVLLHPLGGTWDKDDVPLS 319
QY 457 WRMKQHAALVEEGVLPKSTIVAIFPSPMLYAGTEVQWCHRCRMIAGANFYIVGRDPAG 516
DB 320 WRMKQHEKVLVEEGLNPESTVVAIFPSPMHYAGTEVQWCHRCRMIAGANFYIVGRDPAG 379
QY 517 MHPETKDKLYETHGCKVLSMAPGLTSVEIIFRVAAYNAKAKAMDYDPAHNEFDFFI 576
DB 380 MSHPTEKRDLYDADHGGKVLMSAPGLERLNLFPKVAAYDTKQKMDFFDPSRKDDFLFI 439
QY 577 SGTMRKLAREGENPPDGFMAPKAWKVLTDYYS 610
DB 440 SGTMRKLAREGENPPDGFMAPKAWKVLTDYYS 473

RESULT 11
T08594
probable sulfate adenylyltransferase (EC 2.7.7.4) - rape
C:Species: Brassica napus (rape)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T08594
R:Buchanan-Wollaston, V.; Alnsworth, C.
A:Title: Leaf senescence in Brassica napus: cloning of senescence related genes by subtr
A:Reference number: Z16446; MUID:97260386; PMID:9106506
A:Accession: T08594
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-459 <BUC>
A:Cross-references: EMBL:U68218; NID:g1527218; PIDN:AAB53100.1; PID:g1527219
A:Experimental source: cv. Falcon; senescing leaves
C:Genetics:
A:Gene: LSC680
C:Function:
A:Description: catalyzes the activation of sulfate to adenylylsulfate
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology
C:Keywords: nucleotidyltransferase
F:45-444/Domain: sulfate adenylyltransferase homology <SAT>

Query Match 37.1%; Score 1212.5; DB 2; Length 459;
Best Local Similarity 59.0%; Pred. No. 1.6e-82;
Matches 230; Conservative 63; Mismatches 92; Indels 5; Gaps 5;

QY 226 ELFVPEKNLDHVRRAEAT--LPSLSITKLDLQWQVLSSEGWATPLKGFMRKEYLEQVMHFD 284
DB 55 ELIVDPSRRREKHEAATELPVELTAIDQMWHVLSSEGWASPLOGFMRSEFQLQTLHN 114
QY 285 TL-LDDG-VINMSIPVLVPSAEDKTRLEGCSKFVLHG--GRRVAILRDAEFYEHKKEER 341
DB 115 SLRLDDGSVVNVMSVPIVLPIDDEPKASIGESKRVALVSDGNPVAITLDIEYKHPKEER 174
QY 342 CSRWGTCTTKPHI--KMWESGDWLVGGDLQVLEKIRNWDGLDQYRLTPLELKQKCKEM 400
DB 175 IARTWGTATAPLVVDEAITNAGDWLIGDLEVPKYNKNDGLDRFLSPAELRKELEKR 234

Query Match 17.0%; Score 556; DB 2; Length 202;
Best Local Similarity 51.5%; Pred. No. 4.1e-34;
Matches 104; Conservative 41; Mismatches 53; Indels 4; Gaps 1;

QY 15 STNVVYQAHVHVRNKRQGVVGTGRGCTGCTVLTGLSGAGKTTTISFALEYLVSHAIPCY 74
DB 2 ATNITFHGVSVTKREIKFV----GHPGTMITWGLSASGKSTIACALEQYLQRGVTTY 57
QY 75 SLDDGNVRHGLNRLGFSFGDREENIRRAEVAKLAFADAGLVCTISFISPFKADRENARK 134
DB 58 RLDDGNVRHGLNRLGFSFGDREENIRRAEVAKLAFADAGLVCTISFISPFKADRENARK 117
QY 135 IHESAGLPFFETFDVADPLNICESRDVKGLYKRRAGEIKGFTGIDSDYKPEPTEPVLT 194
DB 118 FHKKDGLPFEVVECPVEAEQDPKGLYKRRAGEIKGFTGIDSDYKPEPTEPVLT 177
QY 195 NLSTVSDCVHVVVELLEQNIV 216
DB 178 HTQSIIECEVEKIVNVLLEKDLI 199

RESULT 13
E96912
adenylylsulfate kinase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C:Accession: E96912
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 193, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: E96912
A:Status: preliminary
A:Molecule type: DNA

adenylylsulfate kinase [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C:Accession: T50101
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Harris, D.
submitted to the EMBL Data Library, February 2000
A:Reference number: Z25037
A:Accession: T50101
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-202 <WOO>
A:Cross-references: EMBL:AL158056; PIDN:CAB76273.1; GSPDB:GN00066; SPDB:SPAC1782.11
A:Experimental source: strain 972h(-); cosmid c1782
C:Genetics:
A:Gene: SPDB:SPAC1782.11
A:Map position: 1
A:Introns: 157/1
C:Superfamily: adenylylsulfate kinase; adenylylsulfate kinase homology

Query Match 17.0%; Score 556; DB 2; Length 202;
Best Local Similarity 51.5%; Pred. No. 4.1e-34;
Matches 104; Conservative 41; Mismatches 53; Indels 4; Gaps 1;

QY 15 STNVVYQAHVHVRNKRQGVVGTGRGCTGCTVLTGLSGAGKTTTISFALEYLVSHAIPCY 74
DB 2 ATNITFHGVSVTKREIKFV----GHPGTMITWGLSASGKSTIACALEQYLQRGVTTY 57
QY 75 SLDDGNVRHGLNRLGFSFGDREENIRRAEVAKLAFADAGLVCTISFISPFKADRENARK 134
DB 58 RLDDGNVRHGLNRLGFSFGDREENIRRAEVAKLAFADAGLVCTISFISPFKADRENARK 117
QY 135 IHESAGLPFFETFDVADPLNICESRDVKGLYKRRAGEIKGFTGIDSDYKPEPTEPVLT 194
DB 118 FHKKDGLPFEVVECPVEAEQDPKGLYKRRAGEIKGFTGIDSDYKPEPTEPVLT 177
QY 195 NLSTVSDCVHVVVELLEQNIV 216
DB 178 HTQSIIECEVEKIVNVLLEKDLI 199

RT sulfurylase/adenosine 5'-phosphosulfate kinase isoform SK2.";
 RL J. Biol. Chem. 274:33306-33312(1999).
 RN [6]

RA SEQUENCE FROM N.A. (ISOFORM B).
 Venkatchalam K.V., Fuda H., Strott C.A.;
 RT "3'-phosphoadenosine 5'-phosphosulfate synthase 2b isoform.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]

RP SEQUENCE FROM N.A. (ISOFORM A).
 RA Tissue-Colon;
 RC Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS
 KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE
 ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE
 GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE
 SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS
 YIELDING 3'-PHOSPHOADENYLYLSULFATE (PAPS). ACTIVATED SULFATE DONOR
 USED BY SULFOTRANSFERASES. IN MAMMALS, PAPS IS THE SOLE SOURCE OF
 SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-
 ACTIVATION PATHWAY. MAY HAVE A IMPORTANT ROLE IN SKELETOGENESIS
 DURING POSTNATAL GROWTH (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: ATP + sulfate = diphosphate + adenylylsulfate.
 CC -1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
 phosphoadenylylsulfate.

CC -1- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO
 THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION
 OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.

CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a (shown here) and b; are
 produced by alternative splicing.

CC -1- TISSUE SPECIFICITY: EXPRESSED IN CARTILAGE.

CC -1- DISEASE: DEFECTS IN PAPSS2 ARE THE CAUSE OF SPONDYLOEPHYSEAL
 DYSPLASIA, PAKISTANI TYPE (SEMD), AN AUTOSOMAL RECESSIVE DISEASE.
 THIS FORM OF SEMD IS CHARACTERIZED BY A DYSPLASIA THAT IS
 PRIMARILY EPIPHYSEAL WITH ONLY MILD METAPHYSEAL ABNORMALITIES.

CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE
 FAMILY.

CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE
 ADENYLYLTRANSFERASE FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; AF091242; AAC64583.1; -
 EMBL; AF074331; AAD38423.1; -
 EMBL; AF313907; AAK00296.1; -
 EMBL; AF160509; AAF40307.2; -
 EMBL; AF160503; AAF40307.2; JOINED.
 EMBL; AF160504; AAF40307.2; JOINED.
 EMBL; AF160505; AAF40307.2; JOINED.
 EMBL; AF160506; AAF40307.2; JOINED.
 EMBL; AF160507; AAF40307.2; JOINED.
 EMBL; AF160508; AAF40307.2; JOINED.
 EMBL; AF173365; AAF12761.1; -
 EMBL; AF150754; AAF20366.2; -
 EMBL; BC009894; AAH09894.1; -
 Genew; HGNC:8604; PAPSS2.
 MIM; 603005; -
 InterPro; IPR002891; APS_kinase.
 InterPro; IPR002650; ATP_sulfurylase.
 Pfam; PF01593; APS_kinase; 1.
 Pfam; PF01747; ATP_sulfurylase; 1.
 ProDom; PD002350; APS_kinase; 1.
 ProDom; PD002381; APS_sulfurylase; 1.
 TIGRams; TIGR00455; apsk; 1.
 Transferase; Nucleotidyltransferase; Kinase; Multifunctional enzyme;
 KW ATP-binding; Multigene family; Alternative splicing.
 DE DOMAIN 1 7210 ADENYLYLSULFATE KINASE.

FT DOMAIN 7211 614 SULFATE ADENYLYLTRANSFERASE.
 FT NP_BIND 49 56 ATP (POTENTIAL).
 FT ACT_SITE 123 123 FORMS THE PHOSPHORINE INTERMEDIATE (BY
 FT SIMILARITY).
 FT SITE 511 515 PP-MOTIF (BY SIMILARITY).
 FT VARSPIC 288 288 D -> DGMALP (IN ISOFORM B).
 FT CONFLICT 166 166 R -> K (IN REF. 2).
 FT CONFLICT 361 361 E -> G (IN REF. 3).
 FT CONFLICT 426 426 R -> C (IN REF. 1).
 FT CONFLICT 567 567 P -> L (IN REF. 2).
 SQ SEQUENCE 614 AA; 69500 MW; 52F4B6D972DDA91E CRC64;

Query Match 98.9%; Score 3236.5; DB 1; Length 614;
 Best Local Similarity 99.3%; Pred. No. 1.2e-225;
 Matches 611; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MSGIKKOKTENQOKSTNVVYQAHVSRNKRQGVVGTGRGPGCTVWLTGSGAGKTTISF 60
 DB 1 MSGIKKOKTENQOKSTNVVYQAHVSRNKRQGVVGTGRGPGCTVWLTGSGAGKTTISF 60
 QY 61 ALLEYLVSHAIPCYSLDGDVNRHGLNRLNGLFSPGDREENIRRIAEVAKLFADAGLVCTTS 120
 DB 61 ALLEYLVSHAIPCYSLDGDVNRHGLNRLNGLFSPGDREENIRRIAEVAKLFADAGLVCTTS 120
 QY 121 FISPFADRENARKIHESAGLPFFEIFVDAPLNICESRDVKGKYKRAGEIKGFTGIDS 180
 DB 121 FISPFADRENARKIHESAGLPFFEIFVDAPLNICESRDVKGKYKRAGEIKGFTGIDS 180
 QY 181 DYKPEPPELVKTNLSTVSDCVHQVVELLQEQNIVPTTIKIDHELFPENKLDHVRAE 240
 DB 181 DYKPEPPELVKTNLSTVSDCVHQVVELLQEQNIVPTTIKIDHELFPENKLDHVRAE 240
 QY 241 AETLPSLSITKLDLQWVQLSEGWATPLKGFMRKEKYLQVMHFDTLDDGVINMSIPIVL 300
 DB 241 AETLPSLSITKLDLQWVQLSEGWATPLKGFMRKEKYLQVMHFDTLDDGVINMSIPIVL 300
 QY 301 PVSAEDKTRLEGCSKFVLAHGGRVAILRDAEYEHKKEKRCRVGTTCTKPHIKVMV 360
 DB 301 PVSAEDKTRLEGCSKFVLAHGGRVAILRDAEYEHKKEKRCRVGTTCTKPHIKVMV 360
 QY 361 ESCDMLVGGDLQVLEKIRWMDGLDQVRLTPLELKQCKKEMNADAVAFQLRNPVNHGHAL 420
 DB 361 ESCDMLVGGDLQVLEKIRWMDGLDQVRLTPLELKQCKKEMNADAVAFQLRNPVNHGHAL 420
 QY 421 LMDQTCRLLERGYKHPVLLLLHPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTTIVAI 480
 DB 421 LMDQTCRLLERGYKHPVLLLLHPLGGWTKDDVPLDWRMKQHAHVLEEGVLDPKSTTIVAI 480
 QY 481 FPSPMYAGPTEVQWHCRSRMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLMSAP 540
 DB 481 FPSPMYAGPTEVQWHCRSRMIAGANFYIVGRDPAGMPHPETKDKLYEPHGGKVLMSAP 540
 QY 541 GLTSVEIIPFRVAAYNKKAKKAMDFYDPARHNEFDISGTRMRKLAREGENPPDGFMAPKA 600
 DB 541 GLTSVEIIPFRVAAYNKKAKKAMDFYDPARHNEFDISGTRMRKLAREGENPPDGFMAPKA 600
 QY 601 WKVLTDDYRSEMDKN 615
 DB 601 WKVLTDDYRS-LEKN 614

RESULT 2

PPS2_MOUSE STANDARD; PRT; 621 AA.
 ID PPS2_MOUSE
 AC O88428; Q92274;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 2 (PAPS
 synthetase 2) (PAPSS 2) (Sulfurylase kinase 2) (SK2) (SK 2)
 DE [includes: sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylyl
 transferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase
 (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)]

NP_BIND	59	66	ATP (POTENTIAL).
FT	133	133	FORMS THE PHOSPHOSERINE INTERMEDIATE (BY SIMILARITY).
FT			
FT	521	525	PP-MOTIF (BY SIMILARITY).
FT	425	425	H->A: LOSS OF ACTIVITY.
FT	426	426	N->K: INCREASED ACTIVITY.
FT	427	427	G->A: 30% DECREASE IN ACTIVITY.
FT	428	428	H->A: LOSS OF ACTIVITY.
FT	427	428	GH->AA: LOSS OF ACTIVITY.
FT	CONFLICT -270	270	L -> F (IN REF. 2).
FT	456	456	MISSING (IN REF. 2 AND 3).
FT	CONFLICT 587	587	S -> L (IN REF. 2).
FT	SEQUENCE 624	AA; 70847	MW; 6BC4F9648016CA31 CRC64;
Query Match			79.4%; Score 2596; DB 1; Length 624;
Best Local Similarity			77.0%; Pred. No. 1.8e-179;
Matches	471; Conservative	67; Mismatches 68; Indels	6; Gaps 1;
QY	5	KKOKTEN-----	QQKSTNNVYQAHVHSRNRKQVVGTRGGFRGCTVWLTGLSGAGKTTT 58
DB	9	KKVLSNNAQNGMRATNTVYQAHVHSRNRKQVVGTRGGFRGCTVWLTGLSGAGKTTV 68	
QY	59	SFAEEVLYSHATPCYSLDGDNVYHGLNRLNGLSPGDREENIRIAEVAKLFDADAGLVCI 118	
DB	69	SMALIEVLYCHGIPCTYLDGDNIRQGLKNLGFSPEDREENVRIAEVAKLFDADAGLVCI 128	
QY	119	TSPISPFADRENARKITHESAGLPFFEIFVDAPLNTICESDRVKGLYKRARAGETKGTGI 178	
DB	129	TSPISPYTDQRNARQIHEGASLPFFEVFVDAPLHVCEQRDVKGLYKKARAGEIKGTGI 188	
QY	179	DSYKPEPTPVRVLKNLSTVSCVHQVVELLOEQNIVPYTIIKDIHELFPVPENKLDHVR 238	
DB	189	DSYEKPEAPELVKLTDCSDVNCVQVVELLOERDIPVPVDASYEVKELYVPENKLDHVR 248	
QY	239	AEATLPSLITKLDLQWQVVLSEGWATPLKGFMEKEYLQVMHFDTLDDGVNKNMSPI 298	
DB	249	TDATLPAULKINKVQWQVVLAEGWATPLNGFMREYRLOCLHFDCLDDGVNLSVPI 308	
QY	299	VLPVSAEDKTRLEGCSKFVLHAGRRVAILRDAAEFVEHRRKEERC SRVWGTCTCKPHIKM 358	
DB	309	VLTAETHEDKRLDGCATAFALWYEGRRVAILRNPEFEHRRKEERCARQWGTCKNHPYIKM 368	
QY	359	VMESGDLWVGGDLQVLEKIRWNDGLQYRLTPIELKQCKEMNADAVFATQLRNPVHNGH 418	
DB	369	VMEQGDWLIGGDLQVLDVYVWNDGLQYRLTPTLQKFKDKMADAVFATQLRNPVHNGH 428	
QY	419	ALLMQDTCRLLERYKHPVLLHPLGGWTKDDVDPLDWKMOHAAVLEEGVLDPKSTIV 478	
DB	429	ALLMQDTHKOLLERYRRPVLLLHP LGAWTKDDVDPLMWKMOHAAVLEEGVLPETTV 488	
QY	479	AIFSPMLVAGPTEVQWCHRSRM IAGANFYIVGRDPAGMPHPTTKDLYEPTHGKVLSM 538	
DB	489	AIFSPMVAAGPTEVQWCHRCARMVAGANFYIVGRDPAGMPHPTTKDLYEPTHGKAVLTM 548	
QY	539	APGLTSVEIIPFVAAVYNAKAKMDFYDPARHNEFDIFSGTRMKRLAREGENPDGFMAP 598	
DB	549	APGLTLEIVPFVAAVYNAKKKRM DYDSHEDDEFISGTRMKRLAREGOKPPEGFMAP 608	
QY	599	KANKVLTDYRS 610	
DB	609	KAWTVLLEYYS 620	
RESULT 5			
PPSL_CAVPO		STANDARD;	PRT; 624 AA.
ID	PPSL_CAVPO		
AC	O54820;		
DT	30-MAY-2000	(Rel. 39, Created)	
DT	30-MAY-2000	(Rel. 39, Last sequence update)	
DT	16-OCT-2001	(Rel. 40, Last annotation update)	
DE	Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPS synthetase 1) (PAPSS 1) (Sulfurylase kinase 1) (SK1) (SK 1)		
DE	[Includes: Sulfate adenylyltransferase (EC 2.7.7.4) (Sulfate adenylyltransferase 1) (PAPS 1) (Sulfurylase kinase 1) (SK1) (SK 1)]		

DE transferase) (SAT) (ATP-sulfurylase); Adenylylsulfate kinase
 DE (EC 2.7.1.25) (Adenylylsulfate 3'-phosphotransferase) (APS kinase)
 DE (Adenosine-5'-phosphosulfate 3'-phosphotransferase) (3'-
 DE phosphoadenosine-5'-phosphosulfate synthetase)).
 GN PAPSS1 OR PAPSS OR ATPSK1
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=NH 2;
 RC Venkatchalam K.V., Akita H., Strott C.A.;
 RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RL -!- FUNCTION: BIFUNCTIONAL ENZYME WITH BOTH ATP SULFURYLASE AND APS
 CC KINASE ACTIVITY, WHICH MEDIATES TWO STEPS IN THE SULFATE
 CC ACTIVATION PATHWAY. THE FIRST STEP IS THE TRANSFER OF A SULFATE
 CC GROUP TO ATP TO YIELD ADENOSINE 5'-PHOSPHOSULFATE (APS), AND THE
 CC SECOND STEP IS THE TRANSFER OF A PHOSPHATE GROUP FROM ATP TO APS
 CC YIELDING 3'-PHOSPHOADENYLYLSULFATE (PAPS: ACTIVATED SULFATE DONOR
 CC USED BY SULFOTRANSFERASE). IN MAMMALS, PAPS IS THE SOLE SOURCE OF
 CC SULFATE; APS APPEARS TO BE ONLY AN INTERMEDIATE IN THE SULFATE-
 CC ACTIVATION PATHWAY. ALSO INVOLVED IN THE BIOSYNTHESIS OF SULFATED
 CC L-SELECTIN LIGANDS IN ENDOTHELIAL CELLS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: ATP + sulfate -> adenylylsulfate.
 CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate -> ADP + 3'-
 CC phosphoadenylylsulfate.
 CC -!- PATHWAY: BELONGS TO THE SULFATE ASSIMILATION PATHWAY THAT LEADS TO
 CC THE BIOSYNTHESIS OF CYSTEINE AND METHIONINE, AND TO THE SULFATION
 CC OF PROTEINS, CARBOHYDRATES, LIPIDS, DRUGS AND XENOBIOTICS.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE APS KINASE
 CC FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE SULFATE
 CC ADENYLYLTRANSFERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF004875; AAC02266.1; --
 DR InterPro: IPR002891; APS_kinase.
 DR InterPro: IPR002650; ATP-sulfurylase.
 DR Pfam: PF01583; APS_kinase; 1.
 DR Pfam: PF01747; ATP-sulfurylase; 1.
 DR ProDom: PD002350; APS_kinase; 1.
 DR ProDom: PD002381; ATP-sulfurylase; 1.
 DR TIGRFAMs: TIGR00455; apsk; 1.
 DR Transferase: Nucleotidyltransferase; Kinase; Multifunctional enzyme;
 KW ATP-binding; Multigene family.
 FT DOMAIN 1 220
 FT SUFFATE ADENYLYLTRANSFERASE.
 FT NP_BIND 221 624
 FT ACT_SITE 59 66
 FT ACT_SITE 133 133
 FT PP-MOTIF (BY SIMILARITY).
 FT SITE 521 525
 FT SEQUENCE 624 AA; 70395 MW; BF7461B4D07F2131 CRC64;
 SQ
 Query Match 78.4%; Score 2563; DB 1; Length 624;
 Best Local Similarity 77.3%; Pred. No. 4.2e-177;
 Matches 462; Conservative 69; Mismatches 67; Indels 0; Gaps 0;
 QY 13 QKSTNVVQAAHVSVRKQGVVGTGGFRGCTVWLTGLSGAGKTTISFALELYLSHAIP 72
 DB 23 QKSTNVVQAAHVSVRKQGVVGTGGFRGCTVWLTGLSGAGKTTISFALELYLSHAIP 72
 QY 73 CYSLDGNVVRHGLNGLGSPGDEENIRRIAEVAKLFADAGLVCTISFISPAKDRNA 132
 DB 83 CYTLDGDNIRGLNGLGSPGDEENIRRIAEVAKLFADAGLVCTISFISPAKDRNA 142

QY 133 RKIHESAGLPFFEIFVDAPLNTICESRDVKGKYLKRRAGEIKGFTGDSYKPEKTPERV 192
 DB 143 RQIHESAGLPFFEIFVDAPLNTICESRDVKGKYLKRRAGEIKGFTGDSYKPEKTPERV 202
 QY 193 KTNLSTVSCVQVQVVELLOEQNIVPTTIKIDHLEFVFPENKLDHVAEATLPSLSITKL 252
 DB 203 KTDACDVNDVQVQVVELLOEQNIVPTTIKIDHLEFVFPENKLDHVAEATLPSLSITKL 262
 QY 253 DLQWQVVLSEGWATPLKGMREKEYLQVNHFTTLLDDGVNNMSIPIVLVPSAEDKVRLEG 312
 DB 263 DMQWQVVLSEGWATPLKGMREKEYLQVNHFTTLLDDGVNNMSIPIVLVPSAEDKVRLEG 322
 QY 313 CSKFVLAHGRRVAILRDAEFYEHKKEERCSRWGTTCTKPHIKMVMESGOWLVGGDIQ 372
 DB 323 CTFAYLIEGRRVAILRNPFEFFHRRKEERCARQWGTCKSHPIYIKVMQGGOWLVGGDIQ 382
 QY 373 VLEKTRWNGDQYRLTPELEKQCKEMNADAVAFQLRNPVHNGHALLMQDTCRELLER 432
 DB 383 VLDRIYNDGLDQYRLTPELEKQCKEMNADAVAFQLRNPVHNGHALLMQDTCRELLER 442
 QY 433 GYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKSTIVAIFFSPMLYAGPTE 492
 DB 443 GYRPRVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKSTIVAIFFSPMLYAGPTE 502
 QY 493 VOWHCRSRMIAGANFYIVGRDPAGPHPTKDLTEPHGGKVLKSNAPGLTSVEIIPFRV 552
 DB 503 VOWHCRSRMIAGANFYIVGRDPAGPHPTKDLTEPHGGKVLKSNAPGLTSVEIIPFRV 562
 QY 553 AAYNKKAKMDYDPARHNEPFICTRMKRLAREGENPPDGMFAPKANKVLTDDYRS 610
 DB 563 AAYNKKAKMDYDPARHNEPFICTRMKRLAREGENPPDGMFAPKANKVLTDDYRS 620
 RESULT 6
 KAPS EMENI STANDARD; PRT; 206 AA.
 AC Q92203;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
 DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-
 DE phosphotransferase).
 GN SD.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=5072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Clarke D.L., Newbert R.W., Turner G.;
 RT "Cloning and characterisation of the APS kinase gene from
 RT Aspergillus nidulans.";
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
 CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate -> ADP + 3'-
 CC phosphoadenylylsulfate.
 CC -!- PATHWAY: SECOND STEP IN THE SULFATE ASSIMILATION PATHWAY THAT
 CC LEADS TO THE BIOSYNTHESIS OF METHIONINE AND CYSTEINE.
 CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y08866; CAA70089.1; --
 DR InterPro: IPR002891; APS_kinase.
 DR Pfam: PF01583; APS_kinase; 1.
 DR ProDom: PD002350; APS_kinase; 1.

DR	TIGR00455; apsk; 1.	
KW	Transferase; Kinase; Cysteine biosynthesis; ATP-binding;	
KW	Phosphorylation.	
FT	NP_BIND 31 ATP (POTENTIAL).	
FT	ACT_SITE 105 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY	
FT		SIMILARITY).
SQ	SEQUENCE 206 AA; 23028 MW; 6AADD483E2BCA1CD CRC64;	
	Query Match 16.2%; Score 529.5; DB 1; Length 206;	
	Best Local Similarity 51.4%; Pred. No. 2.9e-31;	
	Matches 107; Conservative 33; Mismatches 57; Indels 11; Gaps 3;	
QY	15 STNVVYQAHHYSNRKGVGVTGRGFPGCTVMITGLSGAGKTIIISFALRYLVSHPICY 74 ::: :	
Dd	2 ATNITHA - GLTRNQRNKQ --- KGLTIWTGLSSAGKSTIAYELHQLRLGHAY 56 :: :	
QY	75 SLGDNDVRHGNLNLFSPGDREINRRRAEAVALFADAGLVCIITSFISPFADRDNRARK 134 :: :	
Dd	57 RLGDNDVRHGNLDLGFSADRNENIRRAEAVALFADSSIAITSFISPFADRDTARK 116 :: :	
QY	135 IHE-----SAGLPFFEIFVDAPLNICESRDVKGYKKRAGEIKGTGIDSDYEKPETP 188 : :: :	
Dd	117 LHEVPTPNDSGLPVFEVFVDPIEVAEKRPDKLYKKAREGIKEFTGISPYEAPENP 176 :: :	
QY	189 ERYLKTNLSIVSCVHQVVVELLOEQNIV 216 :: :	
Dd	177 EVHKVNVDUPIQEAVKQIIDYLDSKKLL 204 :: :	

RESULT 7

KAPS_YEAST	STANDARD;	PRT;	202 AA.
ID	KAPS_YEAST	STANDARD;	PRT;
AC	Q02196;		
DT	01-JUL-1993 (Rel. 26, Created)		
DT	01-JUL-1993 (Rel. 26, Last sequence update)		
DT	30-MAY-2000 (Rel. 39, Last annotation update)		
DE	Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-		
DE	5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-		
DE	phosphotransferase).		
GN	MET14 OR YKL001C.		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
CC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
OX	NCBI_TaxID=4932;		
[1]	SEQUENCE FROM N.A.		
RN	STRAIN=AB320;		
RC	MEDLINE=91375456; PubMed=1654509;		
RX	Korch C.; Mountain H.A.; Bystrom A.S.;		
RA	"Cloning, nucleotide sequence, and regulation of MEV14, the gene		
RT	encoding the APS kinase of Saccharomyces cerevisiae.";		
RT	Mol. Gen. Genet. 229:96-108(1991).		
RL	[2]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=S288c;		
RC	MEDLINE=93070612; PubMed=1441752;		
FX	Duesterhoeft A., Philippsen P.;		
RA	"DNA sequencing and analysis of a 24.7 kb segment encompassing		
RT	centromere CN11 of Saccharomyces cerevisiae reveals nine previously		
RT	unknown open reading frames.";		
RL	Yeast 8:749-759(1992).		
CC	-1- FUNCTION: ACTYLATES THE SYNTHESIS OF ACTIVATED SULFATE.		
CC	-1- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-		
CC	phosphoadenylylsulfate.		
CC	-1- PATHWAY: SECOND STEP IN THE SULFATE ASSIMILATION PATHWAY THAT		
CC	LEADS TO THE BIOSYNTHESIS OF METHIONINE AND CYSTEINE.		
CC	-1- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.		
CC	----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and/or commercial		

```

DR Pfam; PF01583; APS_kinase; 1.
DR PRODOM; PD002350; APS_kinase; 1.
DR TIGRFAMS; TIGR00455; apsk; 1.
KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
FT NP_BIND 32 39 ATP (POTENTIAL).
FT ACT_SITE 107 107 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
SQ SEQUENCE 211 AA; 23770 MW; 7DDC4BDA867FE7C2 CRC64;
Query Match 16.0%; Score 522.5; DB 1; Length 211;
Best Local Similarity 50.0%; Pred. No. 9.5e-31;
Matches 105; Conservative 35; Mismatches 59; Indels 11; Gaps 3;
QY 15 STNVVYQAHVSRNKRQGVGTGRGCTVWLTGLSGAGKTTISFALBEYLV-SHAIPC 73
DB 2 STNITHASALTSERTELNRQ---RGLTWLTGLSASGKSLAVELEHQLVDRRVHA 57
QY 74 YLSDGONVRHGLNRNLGSPGDREENIRIRIAEVAKLFDAGLVCTISFISPPFAKDRENAR 133
DB 58 YRLDGDGNIRFGLNKGDLGFSADNENIRIRIAEVAKLFDADNSNTAITSFISPYRKDRDTAR 117
QY 134 KIHESA-----GLPFEFLFVDAPLNICESRDVKGLYKKRAGEIKGFTGSDSYKPEP 187
DB 118 QLHEVATPGEETGLPFEVVDVPEVAQORDPKGLYKKRAGEVKEFTGISAPYEAPAN 177
QY 188 PERVLKTNLSTVSDCVHQVVELLQEQNIYP 217
DB 178 PEVHVKNYELPVQDAVKQIIDYLDTKGYLP 207
RESULT 9
CYC2_BACSU
ID CYC2_BACSU STANDARD; PRT; 199 AA.
AC O06735;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE phosphotransferase).
GN YLSZ.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=98015416; PubMed=9353932;
RA Roche B., Autret S., Levine A., Vannier F., Medina N., Seror S.J.;
RT "A Bacillus subtilis chromosome segment at the 100 degrees to 102
RT degrees position encoding 11 membrane proteins.";
RL Microbiology 143:3309-3312(1997).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Hollappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koeter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinot S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,

```

```

RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rochelle E., Roche M., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Taseato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
CC BIOSYNTHETIC PATHWAY.
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y09476; CAAY0655.1; ALT_INIT.
CC EMBL; Z99109; CAB12931.1; -.
CC Subtilist; BG13105; ylsz.
CC InterPro; IPR002891; APS_kinase.
CC Pfam; PF01583; APS_kinase; 1.
CC PRODOM; PD002350; APS_kinase; 1.
CC TIGRFAMS; TIGR00455; apsk; 1.
KW Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation; Complete proteome.
FT NP_BIND 34 41 ATP (BY SIMILARITY).
FT ACT_SITE 108 108 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SEQUENCE 199 AA; 22304 MW; 90EC37FE4B02A123 CRC64;
Query Match 15.3%; Score 500; DB 1; Length 199;
Best Local Similarity 47.0%; Pred. No. 3.7e-29;
Matches 94; Conservative 43; Mismatches 57; Indels 6; Gaps 2;
QY 17 NVVYQAHVSRNKRQGVGTGRGCTVWLTGLSGAGKTTISFALBEYLVSHAIPCYSYL 76
DB 6 NIIHHPAALSKSDRQLS----NGHKSCVLWFTGLSGSKSVLANAYDEKLYRKGIOSYVL 61
QY 77 DGDNVHRHGLNRNLGSPGDREENIRIRIAEVAKLFDAGLVCTISFISPPFAKDRENARKIH 136
DB 62 DGDNTRHGLNKLGLFTGTDRIENIRIRIGEVAKLFVDSGQMLTAFISPPREDMDYRALF 121
QY 137 ESAGLPPFFEIFVDAPLNICESRDVKGLYKKRAGEIKGFTGSDSYKPEPVLKTNL 196
DB 122 PKG--EFFEIIYVKPLHVEQDRPKLYKKARNGEIKHFTGIDSPYEAPLSPDFIIESDQ 179
QY 197 STVSDCVHQVVELLQEQNIY 216
DB 180 TSISDGDGLIINALQNRGII 199
RESULT 10
KAP2_ARATH
ID KAP2_ARATH STANDARD; PRT; 293 AA.
AC O49196;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenylylsulfate kinase 2, chloroplast precursor (EC 2.7.1.25) (APS

```



```
CC phosphadenylylsulfate.
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
CC BIOSYNTHETIC PATHWAY.
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP001512; BAB05208.1; -
CC InterPro; IPR002891; APS_kinase.
CC Pfam; PF01583; APS_kinase; 1.
CC ProDom; PD002350; APS_kinase; 1.
CC TIGRFAMs; TIGR00455; apsk; 1.
CC Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
KW Phosphorylation; Complete proteome.
FT NP_BIND 36 43 ATP (BY SIMILARITY).
FT ACT_SITE 110 110 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
FT SIMILARITY)
SQ SEQUENCE 202 AA; 22658 MW; 6BB234F9242F1FCD CRC64;
Query Match 14.9%; Score 487.5; DB 1; Length 202;
Best Local Similarity 48.8%; Pred. No. 3e-28;
Matches 100; Conservative 31; Mismatches 65; Indels 9; Gaps 3;
QY 9 TENQKSTNVVYQAHVSRNKRQVVGTRGGFGCVWLTGLSGACKTITISFALEELYVS 68
DB 3 TSNQ--PHIVWHEASVSKEEROK----RNRKSCVWFTGLSGSGKSTLANALDKLFE 55
QY 69 HAIPCYSLDGDNRHGLNRLNGLSPGDRENIRRIAEVAKLFADAGLVCTTSPISPFKD 128
DB 56 EGHSHVLDGDNIRHGLNAGLGFSEEDRKNIRIGEVAKLFVDAGVVTAFISFRED 115
QY 129 RENARKIHESAGLPFFEIFVDAPLNICESRDVGLYKRRAGEIKGFTGIDSDYKEPTP 188
DB 116 RDNVRGILDG--EFTIEVVRCPLETCERKDRPKGLYKARSQDIPEFTGTSPPYEPVNP 173
QY 189 ERVLKTNLSTVSCVHVQVVELLQEQ 213
DB 174 ELIIDTDQLAVEAVEKIYALHAQ 198
RESULT 12
ID CYSC_ECOLI STANDARD; PRT; 200 AA.
AC F23846; Q59389; Q59376; P78105;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-
DE phosphotransferase).
GN CYSC OR B2750 OR Z4058 OR ECS3604.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-8.
RX STRAIN=K12;
RC MEDLINE=92268080; PubMed=1316900;
RA Leyh T.S., Vogt T.F., Suo Y.;
RT "The DNA sequence of the sulfate activation locus from Escherichia
RT coli K-12.";
RL J. Biol. Chem. 267:10405-10410(1992).
RN [2]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND ACTIVE SITE.
```

```
RX MEDLINE=93075778; PubMed=1332767;
RA Satishchandran C., Hickman Y.N., Markham G.D.;
RT "Characterization of the phosphorylated enzyme intermediate formed in
RT the adenosine 5'-phosphosulfate kinase reaction.";
RL Biochemistry 31:11684-11688(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamodousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [6]
RP CHARACTERIZATION.
RC MEDLINE=88115389; PubMed=2828368;
RA Leyh T.S., Taylor J.C., Markham G.D.;
RT "The sulfate activation locus of Escherichia coli K12: cloning,
RT genetic, and enzymatic characterization.";
RL J. Biol. Chem. 263:2409-2416(1988).
CC -!- FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
CC -!- CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
CC phosphoadenylylsulfate.
CC -!- PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
CC BIOSYNTHETIC PATHWAY.
CC -!- SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; W74586; AAA23647.1; -
CC EMBL; M86936; AAA23503.1; -
CC EMBL; U29579; AAC69260.1; -
CC EMBL; AE000358; AAC75792.1; -
CC EMBL; AE005502; AAG57857.1; -
CC EMBL; AP002562; BAB37027.1; -
CC PIR; JN0328; JN0328.
CC PIR; A44200; A44200.
CC EcoGene; EG10185; cysC.
CC InterPro; IPR002891; APS_kinase.
CC Pfam; PF01583; APS_kinase; 1.
CC ProDom; PD002350; APS_kinase; 1.
```


DE kinase) (Adenosine-5'-phosphosulfate kinase) (ATP adenosine-5'-
 GN phosphosulfate 3'-phosphotransferase).
 OS Catharanthus roseus (Rosy periwinkle) (Madagascar periwinkle).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;
 OC Vinceae; Catharanthus.
 OX NCBI_TaxID=4058;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schiffmann S., Schwenn J.-D.:
 RT "Isolation of cDNA clones encoding adenosine-5'-phosphosulfate-kinase
 RT (EC2.7.1.25) from Catharanthus roseus and an isoform (akn2) from
 RT Arabidopsis.";
 RL (1) Plant Gene Register PGR98-116.
 CC (-) FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
 CC (-) CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
 CC phosphoadenylylsulfate.
 CC (-) PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
 CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
 CC BIOSYNTHETIC PATHWAY.
 CC (-) SUBCELLULAR LOCATION: Chloroplast (By similarity).
 CC (-) SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF044285; AAC31145.1;
 CC InterPro: IPR002891; APS_kinase.
 CC Pfam: PF01583; APS_kinase; 1.
 CC ProDom: PD002350; APS_kinase; 1.
 CC TIGRFAMs: TIGR00455; apsk; 1.
 CC Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
 CC Phosphorylation; trans peptide; Chloroplast.
 CC TRANSIT ? CHLOROPLAST (POTENTIAL).
 CC CHAIN ? 312 ADENYLYLSULFATE KINASE.
 CC FT NP_BIND 142 149 ATP (POTENTIAL).
 CC FT ACT_SITE 216 216 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
 CC SIMILARITY).
 CC SEQUENCE 312 AA; 33656 MW; 16BD11FB4B1FE27 CRC64;
 CC
 CC Query Match 14.4%; Score 469.5; DB 1; Length 312;
 CC Best Local Similarity 47.7%; Pred. No. 1.1e-26;
 CC Matches 103; Conservative 30; Mismatches 68; Indels 15; Gaps 4;
 QY 3 GIKKQKTENQOKSTNVVYQAHVSRNKRQGVGTGRGCTVWLTLGLSGAGKTTISFAL 62
 Db 100 GKILQTTVGTSTNLHMKCAVEKSERPQQQ----RCGVITWITGLSGSKSTLACAL 155
 QY 63 EYLVSHAI PCYSLDGNVHRHNLNGLFSGDREENIRIAEVAKLFADAGLVCITSFI 122
 Db 156 SRGLHAGKLT YLDDGNVHRHNLNGLSFKAEDRAENIRKIGEVAKLFADAGVICIASLI 215
 QY 123 SPFAKRENAKTHESAGLP---FFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGID 179
 Db 216 SPYKPPDACSRL-----LPEGDFIEVMDVPLKVCARDPKGLYKLARAGKIKGFTGID 270
 QY 180 SDYEKPEPTEPVLKTNL----STVSDCVHQVVELLQE 212
 Db 271 DPVEPPKSEIVLHQKMGCMDCSPCDLADIVISYLEE 306
 RESULT 15
 CISC_VIBCH
 ID CYSC_VIBCH STANDARD; PRT; 215 AA.
 AC Q9KP21;
 DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Adenylylsulfate kinase (EC 2.7.1.25) (APS kinase) (Adenosine-
 DE 5'-phosphosulfate kinase) (ATP adenosine-5'-phosphosulfate 3'-
 DE phosphotransferase).
 GN CYSC OR VC2558.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483 (2000).
 CC (-) FUNCTION: CATALYZES THE SYNTHESIS OF ACTIVATED SULFATE.
 CC (-) CATALYTIC ACTIVITY: ATP + adenylylsulfate = ADP + 3'-
 CC phosphoadenylylsulfate.
 CC (-) PATHWAY: SECOND STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
 CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
 CC BIOSYNTHETIC PATHWAY.
 CC (-) SIMILARITY: BELONGS TO THE APS KINASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE004324; AAF95699.1;
 CC TIGR: VC2558;
 CC InterPro: IPR002891; APS_kinase.
 CC Pfam: PF01583; APS_kinase; 1.
 CC ProDom: PD002350; APS_kinase; 1.
 CC TIGRFAMs: TIGR00455; apsk; 1.
 CC Transferase; Kinase; Cysteine biosynthesis; ATP-binding;
 CC Phosphorylation; Complete proteome.
 CC FT NP_BIND 46 53 ATP (BY SIMILARITY).
 CC FT ACT_SITE 120 120 FORMS THE PHOSPHOSERINE INTERMEDIATE (BY
 CC SIMILARITY).
 CC SEQUENCE 215 AA; 23906 MW; 83B9EE2F295CDDF6 CRC64;
 CC
 CC Query Match 14.2%; Score 463; DB 1; Length 215;
 CC Best Local Similarity 47.6%; Pred. No. 1.9e-26;
 CC Matches 100; Conservative 33; Mismatches 65; Indels 12; Gaps 3;
 QY 10 ENQOKSTNVVYQAHVSRNKRQGVGTGRGCTVWLTLGLSGAGKTTISFALAEYLVS 69
 Db 11 EODAKPENVMHRHVDKQAKR-----ATLKQORPAVLWFTGLSGAGKSTVAGALENRLAAL 66
 QY 70 AIPCYSLDGNVHRHNLNGLFSGDREENIRIAEVAKLFADAGLVCITSFISPAKOR 129
 Db 67 GYHTYLLDDGNVHRHNLNGLFSGDREENIRIRIGELAKLMDAGDLVLTAFISPHRAER 136
 QY 130 ENARKITHESAGLP---FFEIFVDAPLNICESRDVKGLYKRARAGEIKGFTGIDSVEKPE 186
 Db 127 QMVRDL-----LPNGFELEVYNTSLDVCEARDPKGLYKKARAGEIQFTGIDSAYEAPL 181
 QY 187 IPERVLKTNLSTVSDCVHQVVELLQEQNIV 216
 Db 182 NFDIDLPAKEKSVDELVAQCLQALAEHRHII 211

Search completed: May 30, 2003, 12:18:02
Job time : 25 secs

THIS PAGE BLANK (US70)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 12:14:14 ; Search time 88 Seconds
(without alignments)
1439.988 Million cell updates/sec

Title: US-09-898-165B-7

Perfect score: 3271

Sequence: 1 MSGIKKQKTENQKSTNVVY.....MAPKAWKLTIDYRSEMDKN 615

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.rodent.*
- 12: sp.virus.*
- 13: sp.vertibrate.*
- 14: sp.unclassified.*
- 15: sp.rvirus.*
- 16: sp.bacteriap.*
- 17: sp.archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3020.5	92.3	620	11	Q9JK86	Q9JK86 cavia porce
2	2965.5	90.7	621	11	Q9QXS0	Q9QXS0 mus musculus
3	2638.5	80.7	613	13	Q9QXY2	Q9QXY2 fugu rubrip
4	2602	79.5	624	4	Q96TF4	Q96TF4 homo sapien
5	2596	79.4	624	4	Q9P1P9	Q9P1P9 homo sapien
6	2588	79.1	624	4	Q9UE98	Q9UE98 homo sapien
7	2367	72.4	610	5	Q27128	Q27128 urechis cau
8	2233.5	68.3	618	5	Q9NDP8	Q9NDP8 ciona intes
9	2039	62.3	629	5	Q961A8	Q961A8 drosophila
10	2031.5	62.1	630	5	Q9VW48	Q9VW48 drosophila
11	1887	57.7	652	5	Q22501	Q22501 caenorhabdi
12	1351.5	41.3	265	4	Q9UIR2	Q9UIR2 homo sapien
13	1301	39.8	463	10	Q9LIK9	Q9LIK9 arabidopsis
14	1299	39.7	424	10	Q43170	Q43170 solanum tub
15	1282	39.2	465	10	Q42520	Q42520 arabidopsis
16	1282	39.2	465	10	Q96530	Q96530 arabidopsis

17	1282	39.2	465	10	O23324	O23324 arabidopsis
18	1274	38.9	463	10	O43183	O43183 solanum tub
19	1268.5	38.8	476	10	O43870	O43870 arabidopsis
20	1265	38.7	483	10	O96541	O96541 brassica ol
21	1264	38.6	461	10	O92N28	O92N28 brassica ju
22	1260	38.5	483	10	O9SBL0	O9SBL0 brassica ol
23	1259.5	38.5	469	10	O9S7D8	O9S7D8 arabidopsis
24	1257.5	38.4	470	10	O92N29	O92N29 brassica ju
25	1254.5	38.4	476	10	O8RWJ3	O8RWJ3 arabidopsis
26	1252.5	38.3	458	10	O8W1X3	O8W1X3 allium cepa
27	1250	38.2	463	10	O9SE02	O9SE02 arabidopsis
28	1245	38.1	463	10	O42519	O42519 arabidopsis
29	1244.5	38.0	461	10	O9SDP4	O9SDP4 allium cepa
30	1230.5	37.6	489	10	O48888	O48888 zea mays (m
31	1230	37.6	476	10	O92WM0	O92WM0 oryza sativ
32	1212.5	37.1	459	10	O96349	O96349 brassica na
33	1158.5	35.4	336	5	O95P41	O95P41 aedes aegypt
34	1108.5	33.9	461	10	O8SAG1	O8SAG1 glycine max
35	556	17.0	202	3	O9P7G9	O9P7G9 schizosacch
36	551.5	16.9	200	16	O97MT8	O97MT8 clostridium
37	524	16.0	202	3	O9HGF8	O9HGF8 saccharomyc
38	524	16.0	202	3	O9C2Y6	O9C2Y6 saccharomyc
39	508.5	15.5	201	2	O9EY11	O9EY11 klebsiella
40	493.5	15.1	213	16	O8ZBF3	O8ZBF3 yersinia pe
41	481.5	14.7	201	16	O8XF34	O8XF34 salmonella
42	481.5	14.7	635	16	O9A882	O9A882 caulobacter
43	472.5	14.4	208	10	O9SRW7	O9SRW7 arabidopsis
44	460.5	14.1	290	10	O9FJX1	O9FJX1 arabidopsis
45	452	13.8	645	16	O98504	O98504 rhizobium 1

ALIGNMENTS

RESULT 1

ID	Q9JK86	PRELIMINARY;	PRT;	620 AA.
AC	Q9JK86;			
DT	01-OCT-2000 (Tremblrel_15_Created)			
DT	01-OCT-2000 (Tremblrel_15, Last sequence update)			
DT	01-JUN-2002 (Tremblrel_21, Last annotation update)			
DE	Adenosine 5'-phosphosulfate kinase/ATP sulfurylase 2.			
OS	Cavia porcellus (Guinea pig).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.			
OX	NCBI_TaxID=10141;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Fuda H., Strott C.A.;			
RT	"Guinea pig bifunctional adenosine 5'-phosphosulfate kinase/ATP sulfurylase 2, gpAPS synthase 2."			
RL	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF251798; AAF70194.1;			
DR	InterPro; IPR002891; APS_kinase.			
DR	InterPro; IPR002650; ATP_sulfurylase.			
DR	Pfam; PF01583; APS_kinase; 1.			
DR	Pfam; PF01747; ATP_sulfurylase; 1.			
DR	ProDom; PD002350; APS_kinase; 1.			
DR	ProDom; PD002381; APS_kinase; 1.			
DR	TIGRFAMs; TIGR00455; apsk; 1.			
KW	Kinase.			
SQ	SEQUENCE 620 AA: 70031 MW: 50E187D16335985A CRC64;			
Query Match 92.3%; Score 3020.5; DB 11; Length 620;				
Best Local Similarity 91.5%; Pred. No. 9.2e-235;				
Matches 563; Conservative 26; Mismatches 21; Indels 5; Gaps 1;				
QY	1 MSGIKKQKTENQKSTNVVYQAHVSRNRKRGQVGTGRGCTVWLTGLSGAGKTTISF 60			
	: : : : : : : : :			
DB	1 MSGVKKQKTESQKSTNVVYQAHVSRNRKRGQVGTGRGCTVWLTGLSGAGKTTISF 60			
	: : : : : : : : :			
QY	61 ALEEVLVSHAIPCYSLDGDNVPHGLNRNLGFGPDREENIRIAEVAKIFADAGLVCTIS 120			
	: : : : : : : : :			

Db	61	ALB EYLSM SHALPCYSLD GDNVHGLNKLNGFSPGDRENIIRIAEVAKLFA DAGLVCIITS	121
Qy	121	FISPF AKDRENARKIHESAGL PFFEIFYDAPL NICESRDV KGLYKRARAGEIKGFTGIDS	180
Db	121	FISPF TKDRENARKIHESAGL PFFEIFYDAPL NICESRDV KGLYKRARAGEIKGFTGIDS	180
Qy	181	DYK EPTPEPVLKTNLSTVSCVHQVVELLOEQNLVPTIIKDIIHEL FVPENKLDHVRAE	240
Db	181	DYK EPTPECVKLTNLSTVSCVQOVVELLOEQNLPHPTIVKGIHEL FVPENKLDQVRTE	240
Qy	241	AETLPSLSITKLDLQWQVQLSEGWATPLKGFMRKEYLQVWHFDLLD-----DGVINMS	295
Db	241	AESLPSLSITKLDLQWQVQLSEGWATPLKGFMRKEYLQTLHFDFTL DCGVLRDGVINLS	300
Qy	296	IPIVLPVSAEDKTRLEGCSKFVLAHGGRVAILRDAEYF EHRKEBRCRSVGTCTCKHPH	355
Db	301	VPIVLPVSAADKARLEGCSFALMYGGRVAILDCDFEY EHRKVERCCRVNGTSSAKHPH	360
Qy	356	IKMYMESGDLWVGGLDQVLEKIRWN DGLDQVRLTPLELKQCKKENNADAVAFALRNPVH	415
Db	361	VKMYMESGDLWVGGLDQVLERIRWN DGLDKYRLTPLELKQCKKENNADAVAFALRNPVH	420
Qy	416	NGHALLMQDTCRRLLERGYKHPVLLLP LGWTKDDDDVPLDWRMKQHAALVEEGVLDPKS	475
Db	421	NGHALLMQDTHRQLLERYGKHPVLLLP LGWTKDDDDVPLDWRMKQHTAVLEEGVLDPKS	480
Qy	476	TIVAIFSPMLYAGPTEVQWHCRSRMTAGANFYIVGRDPACMPHPETKKDLYEPTHGKV	535
Db	481	TIVAIFSPMLYAGPTEVQWHCRSRMTAGANFYIVGRDPACMPHPETKKDLYEPTHGKV	540
Qy	536	LSMAGLTSVEIIPFRVAAYNKA KAMDIFYDPARHNEFDFISGTRMRKLAREGENPPDGF	595
Db	541	LSMAGLTSVEIIPFRVAAYNKVKAMDIFYNPERHDEFDFISGTRMRKLAREGENPPDGF	600
Qy	596	MAPKAKVLTDIYRS 610	
Db	601	MAPKAKVLTDIYRS 615	
RESULT 2			
Q9QY50			
ID	Q9QY50	PRELIMINARY;	PRT; 621 AA.
AC	Q9QY50		
DT	01-MAY-2000	(T=EMBLrel. 13, Created)	
DT	01-MAY-2000	(T=EMBLrel. 13, Last sequence update)	
DT	01-JUN-2002	(T=EMBLrel. 21, Last annotation update)	
DE	APP sulfurylase/APS kinase isoform sk2 (EC 2.7.7.4).		

RESULT 2

RESULT 2	Q9QYS0	PRELIMINARY;	PRT;	621 AA.
ID	Q9QYS0			
AC	Q9QYS0;			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	ATP sulfurylase/APS kinase isoform SK2 (EC 2.7.7.4).			
DE	PAPSS2.			
GN	Mus musculus (Mouse).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RP	[1]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6;			
RX	MEDLINE=98337975; PubMed=9671738;			
RA	Kurima K., Warman M.L., Krishnan S., Domowicz M., Krueger R.C. Jr.,			
RA	Deypur A., Schwartz N.B.;			
RT	"A member of a family of sulfate-activating enzymes causes murine			
RT	brachymorphism.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 95:8681-8685(1998).			
RN	[2]			
RC	SEQUENCE FROM N.A.			
RP	STRAIN=C57BL/6;			
RX	MEDLINE=20026854; PubMed=10559207;			
RA	Kurima K., Singh B., Schwartz N.B.;			
RA	"Genomic organization of the mouse and human genes encoding the ATP			
RT	Sulfurylase/Adenosine 5'-phosphosulfate kinase isoform SK2.";			
RT	J. Biol. Chem. 274:33306-33312(1999).			
RL	EMBL; AF172865; AAF12760.1; .			
DR	EMBL; AF172857; AAF12760.1; JOINED.			
DR	EMBL; AF172858; AAF12760.1; JOINED.			
DR	EMBL; AF172859; AAF12760.1; JOINED.			

QY 299 VLPVSAEDKTRLEGCSEKFLVLAHGGRRVAILRDAEFYEHRRKEERCSRVWGTCTTKRPHIKM 358
 DB 309 VLTATHEDKERLDGCTAFALMYEGRRVAILRNPPEFFEHRRKEERCARQGTCKNHPYIKM 368
 QY 359 VMESGDNLVGGDLQVLEKIRWNGDGLQYRLTPELKKOKCKEMNADAVAFQLRNPVHNGH 418
 DB 369 VMESGDNLVGGDLQVLEKIRWNGDGLQYRLTPELKKOKCKEMNADAVAFQLRNPVHNGH 428
 QY 419 ALLMODTCRLLRGYKHPVLLHPLGWTGKDDVPLDWRMKQHAALVEEGVLDPKSTIV 478
 DB 429 ALLMODTHKOLLERGYRRPVLLHPLGWTGKDDVPLDWRMKQHAALVEEGVLDPKSTIV 488
 QY 479 AIFPSPMLYAGPTEVQVHCHSRMTAGANFYIVGRDPAGMPHPETGKDYEPHGGKVLISM 538
 DB 489 AIFPSPMLYAGPTEVQVHCHSRMTAGANFYIVGRDPAGMPHPETGKDYEPHGGKVLISM 548
 QY 539 APLGTSVEIIPFVAAAYNKKAKMDFYDPAHNEFDFISGTRMKKLAREGENPPDGFMAP 598
 DB 549 APLGTSVEIIPFVAAAYNKKAKMDFYDPAHNEFDFISGTRMKKLAREGENPPDGFMAP 608
 QY 599 KAWKVLTDYYS 610
 DB 609 KAWTVLTEYYS 620

RESULT 5

Q9P1P9 ID Q9P1P9 PRELIMINARY; PRT; 624 AA.
 AC Q9P1P9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE 3'-phosphoadenosine 5'-phosphosulfate synthetase.
 GN PAPSS.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=20145452; PubMed=10679223;
 RA Xu Z.-H., Otterness D.M., Freimuth R.R., Carlini E.J., Wood T.C.,
 RA Mitchell S., Moon E., Kim U.-J., Xu J.-P., Siciliano M.J.,
 RA Weinshilboum R.M.;
 RT "Human 3'-phosphoadenosine 5'-phosphosulfate synthetase 1 (PAPSS1) and
 RT PAPSS2: gene cloning, characterization and chromosomal localization.";
 RL Biochem. Biophys. Res. Commun. 268:437-444(2000).
 DR EMBL: AF105227; AAF40236.1; -;
 DR InterPro: IPR002891; APS_kinase.
 DR InterPro: IPR002650; ATP-sulfurylase.
 DR Pfam: PF01583; APS_kinase; 1.
 DR Pfam: PF01747; ATP-sulfurylase; 1.
 DR ProDom: PD002350; APS_kinase; 1.
 DR ProDom: PD002381; ATP-sulfurylase; 1.
 DR TIGRfams: TIGR00455; apsk; 1.
 SQ SEQUENCE 624 AA; 70859 MW; 60DC9B943E7B75ED CRC64;

Query Match 79.4%; Score 2596; DB 4; Length 624;
 Best Local Similarity 77.0%; Pred. No. 1.5e-200;
 Matches 471; Conservative 67; Mismatches 68; Indels 6; Gaps 1;

QY 5 KKOKTEN-----QQKSTNVVYQAHVSRNKRQGVVTRGGFRGCTVWLTGLSGAGKTTI 58
 DB 9 KVKVLSNNAQNMGMORATNTVYQAHVSRNKRQGVVTRGGFRGCTVWLTGLSGAGKTTV 68
 QY 59 SFALLEYLVSHATPCYSLDGDVNRHGLNRLNGLGSPGDRENRIRIAEVAKLADAGLVCI 118
 DB 69 SMALEYLVCHGIPCTVLDGDNIRQGLNRLNGLGSPEDRENRIRIAEVAKLADAGLVCI 128
 QY 119 TSFISPFADKRENRARIHESAGLPFFEIFVDAPLNICESRDVKGLYKRRAGEIKGFTGI 178

DB 129 TSFISPYTDQRNNARQIHEGASLPFFEFVDPALHVCQEQRDVKGLYKRRAGEIKGFTGI 188
 QY 179 DSDYEKETPERVLKTNLSTVSDCVHVVWELLOQNTVPYTIKIDIELFLFVPENKLDHVR 238
 DB 189 DSEYKEPEAPELVLTQSCDNDCCVQVVELLQERDIPVDASVEYKELYVPENKHLIAK 248
 QY 239 AEATLPSLSITTKLDLQWVQVLSSEGWATPLKGFMEKREYLVQMHFDTLDDGVINMSIPI 298
 DB 249 TDAETLPALKINKVDMQVQVLAESGWATPLNGFMRERYLOCLHFDCLLDGVLNLSVPI 308
 QY 299 VLPVSAEDKTRLEGCSEKFLVLAHGGRRVAILRDAEFYEHRRKEERCSRVWGTCTTKRPHIKM 358
 DB 309 VLTATHEDKERLDGCTAFALMYEGRRVAILRNPPEFFEHRRKEERCARQGTCKNHPYIKM 368
 QY 359 VMESGDNLVGGDLQVLEKIRWNGDGLQYRLTPELKKOKCKEMNADAVAFQLRNPVHNGH 418
 DB 369 VMESGDNLVGGDLQVLEKIRWNGDGLQYRLTPELKKOKCKEMNADAVAFQLRNPVHNGH 428
 QY 419 ALLMODTCRLLRGYKHPVLLHPLGWTGKDDVPLDWRMKQHAALVEEGVLDPKSTIV 478
 DB 429 ALLMODTHKOLLERGYRRPVLLHPLGWTGKDDVPLDWRMKQHAALVEEGVLDPKSTIV 488
 QY 479 AIFPSPMLYAGPTEVQVHCHSRMTAGANFYIVGRDPAGMPHPETGKDYEPHGGKVLISM 538
 DB 489 AIFPSPMLYAGPTEVQVHCHSRMTAGANFYIVGRDPAGMPHPETGKDYEPHGGKVLISM 548
 QY 539 APLGTSVEIIPFVAAAYNKKAKMDFYDPAHNEFDFISGTRMKKLAREGENPPDGFMAP 598
 DB 549 APLGTSVEIIPFVAAAYNKKAKMDFYDPAHNEFDFISGTRMKKLAREGENPPDGFMAP 608
 QY 599 KAWKVLTDYYS 610
 DB 609 KAWTVLTEYYS 620

RESULT 6

Q9UE98 ID Q9UE98 PRELIMINARY; PRT; 624 AA.
 AC Q9UE98;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE ATP sulfurylase/APS kinase.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RA Deyrup A.T.;
 RA "Human ATP sulfurylase/APS kinase.";
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF016496; AAD09325.1; -;
 DR InterPro: IPR002891; APS_kinase.
 DR InterPro: IPR002650; ATP-sulfurylase.
 DR Pfam: PF01583; APS_kinase; 1.
 DR Pfam: PF01747; ATP-sulfurylase; 1.
 DR ProDom: PD002350; APS_kinase; 1.
 DR ProDom: PD002381; ATP-sulfurylase; 1.
 DR TIGRfams: TIGR00455; apsk; 1.
 KW Kinase.
 SQ SEQUENCE 624 AA; 70881 MW; 0008DBCS7B6F35BC CRC64;

Query Match 79.1%; Score 2588; DB 4; Length 624;
 Best Local Similarity 76.6%; Pred. No. 6.8e-200;
 Matches 469; Conservative 68; Mismatches 69; Indels 6; Gaps 1;

QY 5 KKOKTEN-----QQKSTNVVYQAHVSRNKRQGVVTRGGFRGCTVWLTGLSGAGKTTI 58
 DB 9 KVKVLSNNAQNMGMORATNTVYQAHVSRNKRQGVVTRGGFRGCTVWLTGLSGAGKTTV 68
 QY 59 SFALLEYLVSHATPCYSLDGDVNRHGLNRLNGLGSPGDRENRIRIAEVAKLADAGLVCI 118

[illegible]

```
Query Match      68.3%; Score 2233.5; DB 5; Length 618;
Best Local Similarity 68.2%; Pred. No. 2.5e-171;
Matches 410; Conservative 75; Mismatches 115; Indels 1; Gaps 1;

QY 10 ENQOKSTNVYQAHVSRNKGQVVGTRGFRGCTVWLTGLSGAGKTTISFALEEYLVSH 69
DB 14 EKPMSNIVYQHHVSRDGRQVVG-KGAFRGCTVWLTGLSGAGKSTLSMKLEEYLCCK 72
QY 70 AIPCYSLDGDNVHGLNRNKGSPGDREENIRRIAEVAKLFADAGLVCIITSFSPFAKDR 129
DB 73 GIPAYSLDGNIRHGLNKGDFAPADREENIRRIAEVAKLFADAGLVCIITSFSPYKDR 132
QY 130 ENARKIHESAGLPFFEIFVDAPNICESRDVKGLYKRRAGEIKGFTGIDSDEYKPTPE 189
DB 133 QSARRVHNKSLPFFIEFVDTPLOVCEGRDVGKLYKRRAGEIKGFTGIDSDEYKPTPE 192
QY 190 RVLKTNLSTVSDCVHVVQVVELQEQNIIVPYTIIDHILFVPEKNKLDHVAEATLPSLSI 249
DB 193 VVTKTADMPVEDCMVLVVKELMERNIIPCETTLPVMEFTSHSOVKKRAEALPSIDI 252
QY 250 TKLDQWVQLSEGWATPLKGFREKEYLQVHMFDITLLDDGVNMSIPIVLPVSAEDKTR 309
DB 253 TKLDQWVQLSEGWATPLKGFREKEYLQVHMFDITLLDDGVNMSIPIVLPVSAEDKTR 312
QY 310 LEGCSFVLHAGGRVAILRDAEFYEHKREKRSRVGTTCTKHPIHKVMWESGDMVLVG 369
DB 313 VENAETALYDQVKAHLKPEFYPHLKEERSQWGTSGNKGPHIKMIESGDMVLGCG 372
QY 370 DLQVLEKIRWNGDQVRLTLPLELKOKCEKMNADAVAFOLRNPVNHGALLMODTCRL 429
DB 373 DIEVLRIWGDGLDYKWTPLERAKFKSMNADAVAFOLRNPVNHGALLMODTCRL 432
QY 430 LERKYHVPVLLHPLGGWTKDDVPLDWRMKQHAALVEEGVLPKSTIVAIFFSPMLYAG 489
DB 433 VERGEKNPVLHPLGGWTKDDVPLDWRMKQHAALVEEGVLPKSTIVAIFFSPMLYAG 492
QY 490 PTEVQHCRSRMTAGANFYIVGRDPAGMPHPETKDDLYEPHGHGKVLSPMLYVLIIP 549
DB 493 PTEVQHAKARLATGANFYIVGRDPAGMPHPEDKRDLYDHSHGKVLSPMLYVLIIP 552
QY 550 FRVAAYNKAKAMDYDPPARHNEFDFISGTRMRKLAREGPPDGFMAPKAMKVLVDYR 609
DB 553 FRVAAYNKTKKMDYDPPENHNEFDFISGTRMRKLAREGPPDGFMAPKAMKVLVDYR 612
QY 610 S 610
DB 613 S 613

RESULT 9
Q961A8 PRELIMINARY; PRT; 629 AA.
ID Q961A8;
AC Q961A8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE L225351P.
GN PAPS OR CG8363.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP STRAIN-BERKELEY;
RC Stapleton M., Brokstein P., Hong L., Agbani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY0511724; AAK93148.1; -.
DR FlyBase; FBgn020389; Paps.
DR InterPro; IPR002891; APS_Kinase.
DR InterPro; IPR002650; ATP-sulfonylase.
DR Pfam; PF01593; APS_kinase; 1.
DR Pfam; PF01747; ATP-sulfonylase; 1.
DR ProDom; PD002350; APS_kinase; 1.
DR ProDom; PD002381; ATP-sulfonylase; 1.
DR TIGRFAMs; TIGR00455; aps; 1.
SQ SEQUENCE 629 AA; 71146 MW; 6FF3F8797D98CCB3 CRC64;

Query Match      62.3%; Score 2039; DB 5; Length 629;
Best Local Similarity 63.1%; Pred. No. 1.2e-155;
Matches 392; Conservative 95; Mismatches 120; Indels 14; Gaps 8;

QY 2 SGIKKOKTNOQKSTNVYQAHVSRNKGQVVGTRGFRGCTVWLTGLSGAGKTTISFA 61
DB 7 SSKRKOKT-CLQVATNTVEQHHVTRETRGKNLGCRGCTVWLTGLSGAGKTTISFA 65
QY 62 LEBYLVSHPALPCYSLDGNVHGLNRNKGSPGDREENIRRIAEVAKLFADAGLVCIITSF 121
DB 66 LEAYLVSRGIPAYGLDGNIRTNGLNGLFTPADREENIRRVGEVAKLFADSGVVAICSF 125
QY 122 ISPFADKRENAKIHESAGLPFFEIFVDAPNICESRDVKGLYKRRAGEIKGFTGIDS 181
DB 126 VSPFADKRENAKIHESAGLPFFEIFVDAPNICESRDVKGLYKRRAGEIKGFTGIDS 185
QY 182 YKPEPTEPVERVLKTNLSTVSDCVHVVQVVELQEQNIIVPYTIIDHILFVPEKN 237
DB 186 YKPEPTEPVERVLKTNLSTVSDCVHVVQVVELQEQNIIVPYTIIDHILFVPEKN 244
QY 238 RABAETLPSLSITKLDQWVQLSEGWATPLKGFREKEYLQVHMFDITLLD--DGV--N 293
DB 245 RHEAESLQATEISTVELQWVLAEGWAYPLRGFMREDEYLTQTLHFNLTQSGMDSYREN 304
QY 294 MSTPIVLPVSAEDKTRLEGCSKFLAHGGRVAILRDAEFYEHKREKRSRVGTTCTKH 353
DB 305 HSPVIVLSATQADKDRDGCSSLTLYQGRAVAILRPERFYFORKEERLARQGTSPNH 364
QY 354 PHIKVMWESGDMVLGGDLQVLEKIRWNGDQVRLTLPLELKOKCEKMNADAVAFOLRNP 413
DB 365 PYSKQVYESGDLVVGDLAVIERIRWEDGLDQVRLTPNELRRRPFKELNADAFOLRNP 424
QY 414 VNHGALLMODTCRRLERKYHVPVLLHPLGGWTKDDVPLDWRMKQHAALVEEGVLD 473
DB 425 INHGHALLMODTCRRLERKYHVPVLLHPLGGWTKDDVPLDWRMKQHAALVEEGVLD 484
QY 474 KSTIVAIFFSPMLYAGPTEVQHCRSRMTAGANFYIVGRDPAGMPHP--ETKKD--LYEP 529
DB 485 EDTVLAIFSPMAYAGPTEVQHAKARLATGANFYIVGRDPAGMPHPAKETYPDGNLYDA 544
QY 530 THGGKVLSPMLYAGPTEVQHCRSRMTAGANFYIVGRDPAGMPHPAKETYPDGNLYDA 589
DB 545 THGARVLKMAQGLDSMEILLPFRVAAYDKSASRMFAFFEPKRDKEFEFISGTRMRKLAKTGA 604
QY 590 NPPDGFMAPKAMKVLVDYR 610
DB 605 SPDPGFMEPEAWRILATYQN 625

RESULT 10
Q9VW48 PRELIMINARY; PRT; 630 AA.
ID Q9VW48;
AC Q9VW48;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Bifunctional ATP sulfonylase/APS kinase (EC 2.7.7.4) (Sulfate
DE adenylyltransferase) (Sulfate adenylyl transferase) (ATP-sulfonylase)
DE (PAPS protein) (CG8363 protein).
OS PAPS OR CG8363.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
```

18 QVATNVTEQKHVHVTRETRCKNLGCRGFRGCTVWLITGLSGAGKTSIAFEALYLVSRGIP 77

73 CYSLGDNDVNRHGLNRLNGSPGDREENIRRIAEVAKFLADAGLVCTTSITSPFAKDRENA 132

78 AYLGDNDNIRTGKLNGLGFTPADREENIRRVGEVAKFLADSGVVAICSFVSPFADDEMA 137

133 RKIHESAGLPPFEIKFVDAPLNICESRDVGLYKRAGEIKGFTGIDSDEYKPTETPERVL 192

138 RKIHKGADAGLKFEIFVDTPLDVCETRDVAGLYKKAREGVIKGFTGTQTEYERQPMPELV 197

193 KTNLSVSDCVHQVVELLOEONIVPTIYIKDIH--ELFVPEP-KLDHVRAREAEITLPSLS 248

198 NTHGYTVRSSTOKLVTLLQEGELIPKS-LRVDLLPELYPSESIAATEALRHEAESLQAE 256

249 ITKLDLQWQVLSSEGWATPLKGFPMREKEYLQVNMHFDTLDD--DGVV--NMSIPIVLPSVA 304

257 ISTVELQWQVLAEGWAYPLRGPMREDEVLTLLHENTLQSGMDGSSYRENHSHVFIIVLSATQ 316

305 EKTRLGEGSKFVLANGRRVAILRDAEFYEHKREKSCRVWGTTCTKKPHIKMWESGD 364

317 ADKDRLDGGSSTLYKQYKAVAILRRPEFYQRKEERLARQFQTSNPNHPYSQKVYESGD 376

365 WLVGGLDQVLEKIRWNGDLDQVYLTPLLELKQCKEMNADAVAFQLRNPVHNGHALLMQD 424

377 YLVGGDLAVIERLRWEGDLDQYELTNEURLRRFEKLNADAFQLRNPVHNGHALLMQD 436

425 TCRRLLERGYKHPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKSTIVAIFPSP 484

437 TRQQLLERGFKQPVLLHPLGGWTKDDVPLDMRMKQHAALVEEGVLDPKSTIVAIFPSP 496

485 MLYAGTEPVQWCHRSRMIAGANFYIVGRDPAGMHP--EPKKD--LYEPTHGKVLMSAP 540

497 MMYAGTEPVQWHAHAKRNAGANFYIVGRDPAGMHPAKETYPDGNLYDATHGARVLKMAQ 556

541 GLTSVELIPRVAAYNKKAKMDFYDPARHNEFDLSGTBMRLAREGENPPDGEWAPKA 600

557 GLDSMELIPRVAAYDKSASMAFFEPKRDKEEFISGTAKMTLAKTGASPPDGEWPEPA 616

601 WKVLTDYYS 610

617 WRILATYYQN 626

RESULT 11

Q22501 PRELIMINARY; PRT; 652 AA.

Q22501; 01-NOV-1996 (TEMBLrel. 01, Created)

DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)

DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)

DE T14G10.1 protein.

GN T14G10.1

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OC NCBI_TaxID=6239;

RN [1]

RN SEQUENCE FROM N.A.

RA Wild A.;

RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.

RL [2]

RL SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT "genome sequence of the nematode C.elegans: A platform for investigating biology.";

RL Science 282:2012-2018(1998).

DR EMBL; Z68880; CAA93098.1; .

DR InterPro; IPR002891; APS_kinase.

DR InterPro; IPR002650; ATP-sulfurylase.

DR Pfam; PF01583; APS_kinase; 1.

DR Pfam; PF01747; ATP-sulfurylase; 1.

DR ProDom; PD002350; APS_kinase; 1.

[2]	SEQUENCE FROM N.A..
RP	MEDLINE-20026854; PubMed-10559207;
RX	Kurima K., Singh B., Schwartz N.B.;
RA	"Genomic organization of the mouse and
RT	Sulfonylase/Adenosine 5'-phosphosulfate
RT	kinase isoform SK2.";
RL	J. Biol. Chem. 274:33306-33312(1999).
DR	ENBL; AF173364; AAF12780.1; -.
DR	ENBL; AF173361; AAF12780.1; JOINED.
DR	ENBL; AF173362; AAF12780.1; JOINED.
DR	ENBL; AF173363; AAF12780.1; JOINED.
DR	InterPro: IPR002650; ATP-sulfonylase.
DR	Pfam: PF01747; ATP-sulfonylase; 1.
DR	ProDom: PD002381; ATP-sulfonylase; 1.
KW	Kinase; Nucleotidyltransferase; Transferase.
FT	NON_TER
SQ	SEQUENCE 265 AA; 30142 MW; D75A2374E40C334F CRC64;

Query Match	41.3%	Score	1351.5	DB	4	Length	265
Best Local Similarity	95.5%	Prod. No.	8.7e-101				
Matches	254	Conservative	2	Mismatches	1	Indels	9
Gaps							
Qy	358	MVMESGDLVGGDLQVLEKIRWNDGLDQYRLTPTLELKQKCKEMNADAVFAFOLRNPVHG	417				
Db	1	MVMESGDLVGGDLQVLEKIRWNDGLDQYRLTPTLELKQKCKEMNADAVFAFOLRNPVHG	60				
Qy	418	HALLMQDTCRLLERGVKHPVLLLHPGLGGWTQKDDVPLDWRMKQHAALVEEGVLDPKSTI	477				
Db	61	HALLMQDTRKRLERGVKHPVLLLHPGLGGWTQKDDVPLDWRMKQHAALVEEGVLDPKSTI	120				
Qy	478	VAIFPSPMLYAGPTEVQWHCSRMIAGANFYIVGRDPAGMPHPETKDLIYETHGGKVL	537				
Db	121	VAIFPSPMLYAGPTEVQWHCSRMIAGANFYIVGRDPAGMPHPETKDLIYETHGGKVL	180				
Qy	538	MAPCLTSVEIIPFRAAYNAKKAAMDFFYDPA-----RNEFDFI	589				
Db	181	MAPCLTSVEIIPFRAAYNAKKAAMDFFYDPA-----RNEFDFI	240				
Qy	590	NPPDGFMAPKAWKVLTDYYRSEMDKN	615				
Db	241	NPPDGFMAPKAWKVLTDYYRS-LEKN	265				

RESULT 13
Q9LIK9
ID Q9LIK9 PRELIMINARY; PRT; 463 AA.
AC Q9LIK9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 21, Last annotation update)
DE APP sulfurylase/APPS kinase (ATG322890/F5N5_6) (Putative ATP
DE sulfurylase/APPS kinase) (Putative ATP sulfurylase).
GN F5N5_6 OR AT3G22890, F5N5_6 OR AT3G22890.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA MEDLINE=20363099; PubMed=10907853;
RX Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety pI
RT TAC and BAC clones.";
RL DNA Res. 7:217-221(2000).
RN [3]

RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Full length cDNA of gene F5N5.6/AT3g22890 (GI:11994718).";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Davis R.W., Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP001300; BAB03034.1; -
DR EMBL: AY054169; AAL06830.1; -
DR EMBL: AF370492; AAK3869.1; -
DR EMBL: AY050869; AAK92806.1; -
DR EMBL: AY064648; AAL47359.1; -
DR EMBL: AY091207; AAM14146.1; -
DR InterPro: IPR002650; ATP-sulfurylase.
DR Pfam: PF01747; ATP-sulfurylase; 1.
DR ProDom: PD002381; ATP-sulfurylase; 1.
DR TIGRFAMs: TIGR00339; sopt; 1.
KW Kinase.
SQ SEQUENCE 463 AA; 51459 MW; F05DFD8F86CF25C8 CRC64;
Query Match 39.8%; Score 1301; DB 10; Length 463;
Best Local Similarity 62.7%; Pred. No. 2.4e-96;
Matches 244; Conservative 60; Mismatches 81; Indels 4; Gaps 4;
QY 226 ELFPENKLDHVRAEAEETLPSLITKDLQWQVLSGEGWATPLKGFMRKEYLQVWHFDT 285
DB 60 ELIVEEPKREKKHEAADLPVELTAIDLQWHLVSEGWSPLGFGMRSEFLQTLHFNS 119

QY 286 L-LDDG-VINMSIPVLVPSAEDKTRLEGCSKFVLAHG-GRRAVILRDAEFYEHKRERC 342
DB 120 LRLDDGSSVNNKSVPIVLAIDDEQKARIGESTRFALFNSDGNPVALSDIEIYKHPKEERI 179
QY 343 SRVWGTTCTKPHI-KMYMESGDWLVGDLQVLEKIRWNGDLQVRLTPLEUKQCKEKN 401
DB 180 ARTWGTCTAGPLPYDEAITNAGNWLIGDLEVPKYNDGLDRFLSPAELRKELEKRN 239
QY 402 ADAVFAFQLRNPVNHGHALMODTCRRLLERGYKHPVLLHPLGGLGTTKDDVPLDWRMKQ 461
DB 240 ADAVFAFQLRNPVNHGHALMODTCRRLLERGYKHPVLLHPLGGLGTTKDDVPLDWRMKQ 299
QY 462 HAAVLEEGVLDPKSTIVAFITFSPMLYAGTEVQWCHRSRMAGANFYIVGRDPAGMPHPE 521
DB 300 HEKVLEDGVLDPETTVVSIFPSPMHYAGTEVQWCHRSRMAGANFYIVGRDPAGMPHPE 359
QY 522 TKKOLYETHGKGVLSMAPGLTSVEIIPFVAAYNKKAKMDFDPARHNEFDISGTRM 581
DB 360 EKRDLVDADHGKGVLSMAPGLERLNLPRVAAYDKTQGMFAFFDPSRPQDFLISGTNM 419
QY 582 RKLAREGENPPDGFMAPKAWKVLTDYYS 610
DB 420 RTLAKNNENPPDGFMAPKAWKVLTDYYS 448
ID Q43170 PRELIMINARY; PRT; 424 AA.
AC Q43170;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Sulfate adenylyltransferase (EC 2.7.7.4).
GN STWET3-1.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=DESIREE; TISSUE=LEAF;
RX MEDLINE=95004649; PubMed=7920699;
RT Klonus D., Hoefgen R., Willmitzer L., Riesmeier J.W.;
RT "Isolation and characterization of two cDNAs encoding ATP-
RT sulfurylase from potato by complementation of a yeast mutant.";
RL Plant J. 6:105-112(1994).
DR EMBL: X75041; CAA52953.1; -
DR InterPro: IPR002650; ATP-sulfurylase.
DR Pfam: PF01747; ATP-sulfurylase; 1.
DR ProDom: PD002381; ATP-sulfurylase; 1.
DR TIGRFAMs: TIGR00339; sopt; 1.
KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 424 AA; 47519 MW; E82A27DC11350ABC CRC64;
Query Match 39.7%; Score 1299; DB 10; Length 424;
Best Local Similarity 61.7%; Pred. No. 3e-96;
Matches 240; Conservative 64; Mismatches 81; Indels 4; Gaps 4;
QY 226 ELFPENKLDHVRAEAEETLPSLITKDLQWQVLSGEGWATPLKGFMRKEYLQVWHFDT 285
DB 21 ELFVDESQRLKRAEALPKIKTKIDVWVHLSEGWSPLGFGMRSEFLQTLHFNS 80
QY 286 L-LDDG-VINMSIPVLVPSAEDKTRLEGCSKFVLAHG-GRRAVILRDAEFYEHKRERC 342
DB 81 TRLEDGSSVNNKSVPIVLAIDDSQKQIDGSSVALVVDGNPNPAILTDIEYKHPKEERI 140
QY 343 SRVWGTTCTKPHI-KMYMESGDWLVGDLQVLEKIRWNGDLQVRLTPLEUKQCKEKN 401
DB 141 ARTWGTCTAGPLPYDEAITNAGNWLIGDLEVPKYNDGLDRFLSPAELRKELEKRN 200
QY 402 ADAVFAFQLRNPVNHGHALMODTCRRLLERGYKHPVLLHPLGGLGTTKDDVPLDWRMKQ 461

```
||||| 201 ADAVAFQURPNVHNGHALLMTDTRRLLEMGYNKPNVLLHPLGGYTKEDDVPPLHWRMKQ 260
| 462 HAAVLEEGVLDPKSTIVAIFPPSPMLYAGTEVQWCHCRSRMIAGANFYIVGRDPAGMPHPE 521
| 261 HEKVLGGVLDPEITVLSIFPPSPMHYAGTEVQWCHCRSRMIAGANFYIVGRDPAGMSHPL 320
| 522 TKKDLYEPTHGKVLKSNAPGLTSVEIIPRVAAYNKKAKMDFYDPARHNEFDIFSGTRM 581
| 321 EKRDLYDADHGKVLKSNAPGLTSVEIIPRVAAYNKKAKMDFYDPARHNEFDIFSGTRM 380
| 582 RKLAREGENPPDGFMAPKAWKVLTDYYS 610
| 381 RLAKTKESPDPDGFMCPCGKWLVEYYDS 409
```

RESULT 15

```
Q42520 PRELIMINARY; PRT; 465 AA.
AC Q42520;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ATP sulfurylase precursor (EC 2.7.7.4).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96019964; PubMed=7487067;
RA Murillo M., Leustek T.;
RT "Adenosine-5'-triphosphate-sulfurylase from Arabidopsis thaliana and
RT Escherichia coli are functionally equivalent but structurally and
RT kinetically divergent: nucleotide sequence of two adenosine-5'-
RT triphosphate-sulfurylase cDNAs from Arabidopsis thaliana and analysis
RT of a recombinant enzyme.";
RL Arch. Biochem. Biophys. 323:195-204 (1995).
DR EMBL; U06275; AAA92350.1; -
DR InterPro; IPR002650; ATP-sulfurylase.
DR Pfam; PF01747; ATP-sulfurylase; 1.
DR ProDom; PD002381; ATP-sulfurylase; 1.
DR TIGRFAMs; TIGR00339; sotp; 1.
KW Chloroplast; Nucleotidyltransferase; Transferase; Transit peptide.
FT TRANSIT 1 50 CHLOROPLAST.
FT CHAIN 51 465 POTENTIAL.
SQ SEQUENCE 465 AA; 52029 MW; 4405F3E2AEF39735 CRC64;
```

```
Query Match 39.2%; Score 1282; DB 10; Length 465;
Best Local Similarity 62.3%; Pred. No. 8.1e-95;
Matches 243; Conservative 60; Mismatches 81; Indels 6; Gaps 15;

QY 226 ELFVPEKLDHVRARAEATLPSLSITKLDQVQLVSEGWATPLKGFMRKEYLVQVHFD 285
| 62 DLVPEPRREKKHEADLPRLVLTALDQVHVLSEGWASPLRGFMRESEFLQTLHFN 121
| 286 L-LDQG-VINMSIPIVPSAEDKTRLEGCSKFV--LAHGGRVAILRDAEYEHKKEER 341
| 122 LNLDDGSVVNMSVPIVLAIDDOQA-LIGESKRVSLVSDSDNPAILNDIEIYKHPKEER 180
| 342 CSRVWGTTCTKPHI-KWYMESGDWLVGGDLQVLEKIRWNDGLDQYRLTPLEKCKREM 400
| 181 IARTWGTTPAGLPFYVEEAITNAGDWLVGGDLVLEPVYKNDGLRFLRSPFLRKELEKR 240
| 401 NADAVAFQURPNVHNGHALLMTDTRRLLEMGYNKPNVLLHPLGGYTKEDDVPPLDWRMK 460
| 241 GADAVAFQURPNVHNGHALLMTDTRRLLEMGYNKPNVLLHPLGGYTKEDDVPPLDWRMK 300
| 461 QHAAVLEEGVLDPKSTIVAIFPPSPMLYAGTEVQWCHCRSRMIAGANFYIVGRDPAGMPHP 520
| 301 QHEKVLGGVLDPEITVLSIFPPSPMHYAGTEVQWCHCRSRMIAGANFYIVGRDPAGMPHP 360
```

```
QY 521 ETKKDLYEPTHGKVLKSNAPGLTSVEIIPRVAAYNKKAKMDFYDPARHNEFDIFSGTR 580
| 361 VEKRDLYDADHGKVLKSNAPGLTSVEIIPRVAAYNKKAKMDFYDPARHNEFDIFSGTK 420
| 581 MRKLAREGENPPDGFMAPKAWKVLTDYYS 610
| 421 MRALAKNRENPPDGFMCPCGKWLVEYYDS 450
```

Search completed: May 30, 2003, 12:19:39

Job time : 91 secs

Pred. No. is the number of results predicted by chance to have a

TITLE		Mutations in orthologous genes in human spondyloepimetaphyseal dysplasia and the brachymorphic mouse									
JOURNAL	Nat. Genet. 20 (2), 157-162 (1998)										
MEDLINE	98347251										
PUBMED	9771708										
REFERENCE	2 (bases 1 to 2014)										
AUTHORS	ul Haque, M.F., King, L.M., Krakow, D., Cantor, R.M., Rusiniak, M.E., Swank, R.T., Superti-Furga, A., Haque, S., Abbas, H., Ahmad, W., Ahmad, M. and Cohn, D.H.										
TITLE	Direct Submission										
JOURNAL	Submitted (08-SEP-1998) Pediatrics, Medical Genetics, Cedars-Sinai Research Institute, 8700 Beverly Blvd., Los Angeles, CA 90048, USA										
FEATURES	Location/Qualifiers										
source	1..2014										
	/organism="Homo sapiens"										
	/db_xref="taxon:9606"										
	/chromosome="10"										
	/map="10q24"										
	/tissue_type="cartilage"										
	81..1925										
	/note="ATPSK2"										
	/codon_start=1										
	/product="ATP sulfurylase/APS kinase 2"										
	/protein_id="AAC64583.1"										
	/db_xref="GI:3769610"										
	/translation="MSGIKKQKTEENQKSTNVVYQAHVSRNKRQVQVTRGRGRCGT VWLTGSGAGKTTISFALEYLVSHAIPLCYSLDQVNRHGLNRLNGSPDREINIRRR IAEVLADAGLVCTISFPAKRENAKTHESAGLEFFEIFVDAPLNIRESRV KGLYKARAKEIGFTGIDSDYKPEKPERVLKNTLSVSDCHVQVVELLQEQNIQVY YLIKIDHELPENKLDHRAEATPLSLITKLDQWVOVLSEGWATPLKGMREKE TLOVMHFDITLDGVNMSIPIVLPVSAEDKTRLEGSKFLAHGGRVAILRDAEY EHRKERCSRVMTCTCKPHIKVMESGDMVLGCDLOVLEKTRWNGDLQYRLTLE LKQCKEMNADAVFQLRNPVINGHALLMQDRCRLLEKGYKHPVLLHPLPGWTKD DQVPLDWRNKQHAALVEEGVLDPKSTIVAFPSPLMYAGPTEVQVHCRSRMIAGNPF IVGRDPAGMPEKPKDKLDEYTHGKVLSPAGLTISVEIIFRVAAYNKAKKAMDFD PARHNEFDFTSGTRMRKLARENGPPDGFMAKWKVLTIDYRSLEKN"										
	1504										
variation	/note="in SEMD Pakistani type"										
	/replace="a"										
BASE COUNT	515	a	492	c	533	g	474	t			
ORIGIN											
Query Match	100.0%; Score 1845; DB 9; Length 2014;										
Best Local Similarity	100.0%; Pred. No. 0;										
Matches 1845; Conservative	0; Mismatches 0; Indels 0; Gaps 0;										
QY	1	ATGTCGGGGATCAAGAGCAAAAGACGGAGAACACGACAGAAATCCACCAATGTAGTCTAT	60								
DB	81	ATGTCGGGGATCAAGAGCAAAAGACGGAGAACACGACAGAAATCCACCAATGTAGTCTAT	140								
QY	61	CAGGCCACCATCTGACAGCAATAGACAGAGGCAAGTGGTGGACAGGGGTGGGTTTC	120								
DB	141	CAGGCCACCATCTGACAGCAATAGACAGAGGCAAGTGGTGGACAGGGGTGGGTTTC	200								
QY	121	CGAGGATGTACCGTGGGTCTACAGGTCTCTCTGGTGTCTGGAACCAACGATAAGTTTT	180								
DB	201	CGAGGATGTACCGTGGGTCTACAGGTCTCTCTGGTGTCTGGAACCAACGATAAGTTTT	260								
QY	181	GCCTCGGAGGATACCTCTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAT	240								
DB	261	GCCTCGGAGGATACCTCTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAT	320								
QY	241	GTCCGTCATGGCCTTAACAGAAATCTCGGATTCCTCTCTGGGACAGAGAGGAAATATC	300								
DB	321	GTCCGTCATGGCCTTAACAGAAATCTCGGATTCCTCTCTGGGACAGAGAGGAAATATC	380								
QY	301	CGCCGGATGTGAGTGGCTTAAGCTGTTTGTGTGATGCTGGTGTCTGCTGATACCAAGC	360								
DB	381	CGCCGGATGTGCTGAGTGGCTTAAGCTGTTTGTGTGATGCTGGTGTCTGCTGATACCAAGC	440								
QY	361	TTTATTTCTCCATTGCAAGGATCGTGAGATGCGCCGCAAAATACATGAATCAGCAGGG	420								
DB	441	TTTATTTCTCCATTGCAAGGATCGTGAGATGCGCCGCAAAATACATGAATCAGCAGGG	500								

QY	421	CTGCCATTCTTTGAAATATTTGTAGATGCACCTCTAAATATTTGTGAAAGCAGACAGCTA	480
DB	501	CTGCCATTCTTTGAAATATTTGTAGATGCACCTCTAAATATTTGTGAAAGCAGACAGCTA	560
QY	481	AAAGCCCTCTATAAAGAGCCAGAGCTGGGAGATTAAGAGTATTACAGGTATTGATCT	540
DB	561	AAAGCCCTCTATAAAGAGCCAGAGCTGGGAGATTAAGAGTATTACAGGTATTGATCT	620
QY	541	GATTATGAGAAACCTGAAACCTCCTGAGCTGTCTTAAACCAATTTTCCACAGTGGT	600
DB	621	GATTATGAGAAACCTGAAACCTCCTGAGCTGTCTTAAACCAATTTTCCACAGTGGT	680
QY	601	GACTGTGTCTCCACAGTGTGAACTTCTGCAAGAGCAGAACTTGTACCTCTATATA	660
DB	681	GACTGTGTCTCCACAGTGTGAACTTCTGCAAGAGCAGAACTTGTACCTCTATATA	740
QY	561	ATCAAGATATCCACGAACTCTTTGTGCGGAAACAACTTGACCACTCCGAGCTGAG	720
DB	741	ATCAAGATATCCACGAACTCTTTGTGCGGAAACAACTTGACCACTCCGAGCTGAG	800
QY	721	GCTGAACTCTCCCTTCATTATCAATTAAGCTGATCTCCAGTGGTCCAGTCTTTG	780
DB	801	GCTGAACTCTCCCTTCATTATCAATTAAGCTGATCTCCAGTGGTCCAGTCTTTG	860
QY	781	AGCGAAGGCTGGGCACTCCCTCAAGTTTTCATGCGGAGAGAGGAGTACTTACAGTT	840
DB	861	AGCGAAGGCTGGGCACTCCCTCAAGTTTTCATGCGGAGAGAGGAGTACTTACAGTT	920
QY	841	ATGCACCTTGCACACCTCTAGATGATGGCTGATCAACATGACATGACATCCCTTGTACTG	900
DB	921	ATGCACCTTGCACACCTCTAGATGATGGCTGATCAACATGACATGACATCCCTTGTACTG	980
QY	901	CCGCTCTCTGCAGAGGATTAAGACACGCTGGAAGGGTGCACCAAGTTTGTCTGGCAGAT	960
DB	981	CCGCTCTCTGCAGAGGATTAAGACACGCTGGAAGGGTGCACCAAGTTTGTCTGGCAGAT	1040
QY	961	GGTGGAGGAGGAGTGTATCTTACGAGAGCTGAAATCTTATGAACACAGAGAGAGAA	1020
DB	1041	GGTGGAGGAGGAGTGTATCTTACGAGAGCTGAAATCTTATGAACACAGAGAGAGAA	1100
QY	1021	CGCTGTTCCCGTGTTCGGGACAACTGTACAAACACCCCATATCAAAATGCTGATG	1080
DB	1101	CGCTGTTCCCGTGTTCGGGACAACTGTACAAACACCCCATATCAAAATGCTGATG	1160
QY	1081	GAAGTGGGACCTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAGAAAATAAGATGGAAT	1140
DB	1161	GAAGTGGGACCTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAGAAAATAAGATGGAAT	1220
QY	1141	GATGGGCTGACCAATACCGTCTGACACCTCTGAGAGCTCAACAGAAATGTAAGAAATG	1200
DB	1221	GATGGGCTGACCAATACCGTCTGACACCTCTGAGAGCTCAACAGAAATGTAAGAAATG	1280
QY	1201	AATGCTGATCGGTGTTGATTCAGTTCGCAATCTGTCACCAATGGCCATGCCCTG	1260
DB	1281	AATGCTGATCGGTGTTGATTCAGTTCGCAATCTGTCACCAATGGCCATGCCCTG	1340
QY	1261	TTGATGAGGACACCTGCGCCAGGCTCTTAGAGAGGGGCTTACAGCACCCTGCTCCTA	1320
DB	1341	TTGATGAGGACACCTGCGCCAGGCTCTTAGAGAGGGGCTTACAGCACCCTGCTCCTA	1400
QY	1321	CTACACCTCTGGGCGCTGACCAAGGATGACGATGTCCTTAGACTGGCGGATGAAG	1380
DB	1401	CTACACCTCTGGGCGCTGACCAAGGATGACGATGTCCTTAGACTGGCGGATGAAG	1460
QY	1381	CACGACGCGCTGTGCTCGAGGAGGGTCTTGATCCCAAGTCAACATTTGTGGCATC	1440
DB	1461	CACGACGCGCTGTGCTCGAGGAGGGTCTTGATCCCAAGTCAACATTTGTGGCATC	1520
QY	1441	TTTCCGCTCTCCCATGTTATATGTCGCCCCACAGAGGTCCAGTGGGACTGCAGTCCCGG	1500
DB	1521	TTTCCGCTCTCCCATGTTATATGTCGCCCCACAGAGGTCCAGTGGGACTGCAGTCCCGG	1580

```

QY 1501 ATGATTGGGGTGCACATTTCTACATTTGGGGAGGAGCCCTGCGAGGAATGCCCATCT 1560
Db 1581 ATGATTGGGGTGCACATTTCTACATTTGGGGAGGAGCCCTGCGAGGAATGCCCATCT 1640
QY 1561 GAAACCAAGAGGATCTGATGAACCCACTCATGGGGCAAGGCTTGAGCATGCCCT 1620
Db 1641 GAAACCAAGAGGATCTGATGAACCCACTCATGGGGCAAGGCTTGAGCATGCCCT 1700
QY 1621 GGCCTCACCTCTGGAATCATTCATTCGAGTGGCTGCTACAAAGGCAAAAA 1680
Db 1701 GGCCTCACCTCTGGAATCATTCATTCGAGTGGCTGCTACAAAGGCAAAAA 1760
QY 1681 GCCATGACTTCTATGATCCAGCAAGGACAAATGATTTGATCTCAGGAATCGA 1740
Db 1761 GCCATGACTTCTATGATCCAGCAAGGACAAATGATTTGATCTCAGGAATCGA 1820
QY 1741 ATGAGGAAGCTCGCCGGGAAGGAGAGAAATCCCGAGATGGCTTCATGGCCCAAGCA 1800
Db 1821 ATGAGGAAGCTCGCCGGGAAGGAGAGAAATCCCGAGATGGCTTCATGGCCCAAGCA 1880
QY 1801 TGGAGGTCCTGACAGATTTATACAGGTCCTCGGAGAGAACTAA 1845
Db 1881 TGGAGGTCCTGACAGATTTATACAGGTCCTCGGAGAGAACTAA 1925

```

```

RESULT 2
BC009894 2424 bp mRNA linear PRI 12-JUL-2001
LOCUS Homo sapiens, clone MGC:1916 IMAGE:2988345, mRNA, complete cds.
DEFINITION BC009894
ACCESSION BC009894
VERSION BC009894.1 GI:14602765
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2424)
Strausberg,R.
Direct Submission
Submitted (02-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabp3-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fabey, Stephanie Ford, Julia
Greene, Mark Ketteman and Anuradha Madan

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 3 Row: p Column: 17

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 3769609.

```

FEATURES
    Location/Qualifiers
        1..2424
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="MGC:1916 IMAGE:2988345"
            /tissue_type="Colon, adenocarcinoma"
            /clone_lib="NIH_MGC_15"
            /lab_host="DH10B-R"
            /note="Vector: pOTB7"
            121..1965
            /codon_start=1
            /product="unknown (protein for MGC:1916)"
            /protein_id="AAH09894.1"

```

```

/db_xref="GI:14602766"
/translation="MSGIKKQKNTNVYQAHVHVRNKRKQVVGTRGFRGCT
VWLTGSLGAGKTITISFALEYLVSHAIPCYSLIDGNVRHGLNRNLGSPGORENIR
IAEVKLFADAGLVCTISFISPAKEDNRARKIHESAGLPEFIEVDLAPLONIVPY
KGLYKRAAGRIKGTFTGDSYKPEKPELVKLTNLSTVDCVHOVVELLONIVPY
TLIKDIHELFPENKLDHVAEATLPSLITKLDLQVVLSEGAWPLKGFMEKE
YLVQVHFDLLDGGTIVNMSIFIVLPVSAEDKRLSGSKFVLAHGRVRLDRAEFY
EHRKEERSRVWGTCTTKHPIKMNESGDLQVLEKIRWNGDLDQJRLUPLFLE
LKQCKKENNADAVAFQRLNPNVHGHALLMDQTRRLRLRGYKHPVLLHLLPLGGTKD
DVFLDMRKQHAALVEBGLDPKSTIVAFISPMLYAGTEVQWHCSRIMTAGANFY
IVGRDPACMPHETKDIYEPTHGKVLSPAGLTSVEIIPRVAAYNKKAKMDFDYD
PARNERDFISGTRMKRLAREGENPDDFMAPKWKVLDYVRSLEKN"

BASE COUNT      661 a 565 c 596 g 602 t
ORIGIN
Query Match      99.8%; Score 1841.8; DB 9; Length 2424;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1843; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGTCGGGGATCAAGAAGCAAGAGACGAGAGAACCAAGCAATAGTCTAT 60
Db 121 ATGTCGGGGATCAAGAAGCAAGAGACGAGAGAACCAAGCAATAGTCTAT 180
QY 61 CAGGCCCAACCATGTGAGCAGGAATAAGAGAGGCGCAAGTGGTGGACAAAGGGTGGTTC 120
Db 181 CAGGCCCAACCATGTGAGCAGGAATAAGAGAGGCGCAAGTGGTGGACAAAGGGTGGTTC 240
QY 121 CGAGGATGTACCGTGTGCTAACAGAGTCTCTCGTGTCTGGAAAAACACGATAAGTTT 180
Db 241 CGAGGATGTACCGTGTGCTAACAGAGTCTCTCGTGTCTGGAAAAACACGATAAGTTT 300
QY 181 GCCCTGGAGGAGTACCTTGTCTCCCATGCGCATCCCTTGTACTCCCTGGATGGGACAAT 240
Db 301 GCCCTGGAGGAGTACCTTGTCTCCCATGCGCATCCCTTGTACTCCCTGGATGGGACAAT 360
QY 241 GTCCCTCATGGCTTAACAGAAATCTCGGATCTCTCTCGGAGAGAGAGAGAAATATC 300
Db 361 GTCCCTCATGGCTTAACAGAAATCTCGGATCTCTCTCGGAGAGAGAGAGAAATATC 420
QY 301 CCGCGGATGTCTGAGTGGCTAAGCTGTTTGTGTGATGTCTGTCTGTCTGTCTGTCTGTCT 360
Db 421 CCGCGGATGTCTGAGTGGCTAAGCTGTTTGTGTGATGTCTGTCTGTCTGTCTGTCTGTCT 480
QY 361 TTTATTTCTCCATTCGCAAGGATGCTGAGATGCCCCGCAAAATACATCAATCAGCAGG 420
Db 481 TTTATTTCTCCATTCGCAAGGATGCTGAGATGCCCCGCAAAATACATCAATCAGCAGG 540
QY 421 CTGCCATCTTTTGAATATTTGTAGATGCACCTCTAAATATTTTGTGAAAAGCAGACGTA 480
Db 541 CTGCCATCTTTTGAATATTTGTAGATGCACCTCTAAATATTTTGTGAAAAGCAGACGTA 600
QY 481 AAAGGCCCTCTATAAAAGGCCAGAGCTGGGGAGATTTAAAGGATTTACAGTATTGATCT 540
Db 601 AAAGGCCCTCTATAAAAGGCCAGAGCTGGGGAGATTTAAAGGATTTACAGTATTGATCT 660
QY 541 GATTATGAGAAACCTGAAACTCCTGAGCGTGTCTTTAAACCAATTTGTCCAGTGTAGT 600
Db 661 GATTATGAGAAACCTGAAACTCCTGAGCGTGTCTTTAAACCAATTTGTCCAGTGTAGT 720
QY 601 GACTGTGTCACACGAGTACTGGAACCTTCTGCAAGAGCAGCAACATCTTACCTTACTATA 660
Db 721 GACTGTGTCACACGAGTACTGGAACCTTCTGCAAGAGCAGCAACATCTTACCTTACTATA 780
QY 661 ATCAAGATATCCAGGAATCTTTGTGCGGGAACCAACTTTGACCACCTCCGAGCTGAG 720
Db 781 ATCAAGATATCCAGGAATCTTTGTGCGGGAACCAACTTTGACCACCTCCGAGCTGAG 840
QY 721 GCTGAACCTCTCCCTTCAATTAATTAAGTGGATCTCCAGTGGGTCCAGTTTGTG 780
Db 841 GCTGAACCTCTCCCTTCAATTAATTAAGTGGATCTCCAGTGGGTCCAGTTTGTG 900
QY 781 ACAGGAGCTGGCCACCTCCCTCAAGGCTTTCATCGGGGAGAGAGTACTTACAGGTT 840

```

Db 901 AGCGAGGCTGGGCCACTCCCTCAAGGTTTATGCGGGAGAGGAGTACTTACAGGTT 960
QY 841 ATGCACATTTGACACCCCTGCTAGATGATGGGTGATCAACATGAGCATCCCATGTACTG 900
Db 961 ATGCACATTTGACACCCCTGCTAGATGATGGGTGATCAACATGAGCATCCCATGTACTG 1020
QY 901 CCCGTCTCTCAGAGGATAGACACCGCTGGAGGGTGACGAAAGTTTCTTGGGCACAT 960
Db 1021 CCCGTCTCTCAGAGGATAGACACCGCTGGAGGGTGACGAAAGTTTCTTGGGCACAT 1080
QY 961 GGTGGAGGAGGTAGTCTTACGAGACGCTGAATTTCTATGAACACACAAAGAGGAA 1020
Db 1081 GGTGGAGGAGGTAGTCTTACGAGACGCTGAATTTCTATGAACACACAAAGAGGAA 1140
QY 1021 CGCTGTCTCCGTGTTGGGGGACACATGTACAAACACCCCATATCAAAATGGTGATG 1080
Db 1141 CGCTGTCTCCGTGTTGGGGGACACATGTACAAACACCCCATATCAAAATGGTGATG 1200
QY 1081 GAAAGTGGGACCTGGCTGGTGTGGAGACCTTCAGTGTCTGGAGAAATTAAGTGAAT 1140
Db 1201 GAAAGTGGGACCTGGCTGGTGTGGAGACCTTCAGTGTCTGGAGAAATTAAGTGAAT 1260
QY 1141 GATGGGCTGGACCAATACCGCTGTGACACCTCTGGAGCTCAACAGAAATGTAAGAAATG 1200
Db 1261 GATGGGCTGGACCAATACCGCTGTGACACCTCTGGAGCTCAACAGAAATGTAAGAAATG 1320
QY 1201 AATGCTGATGCGGTGTTGCAATTCAGTTCGCAATCTCTCCACATGGCCATGCCCTG 1260
Db 1321 AATGCTGATGCGGTGTTGCAATTCAGTTCGCAATCTCTCCACATGGCCATGCCCTG 1380
QY 1261 TTGATGAGGACACCTGGCCGAGGCTCTAGAGAGGGGTACAGACACCGGTCCTCCTA 1320
Db 1381 TTGATGAGGACACCTGGCCGAGGCTCTAGAGAGGGGTACAGACACCGGTCCTCCTA 1440
QY 1321 CTACACCTCTGGCGGCTGGACCAAGGATGACATGTGCTCTAGACTGGCGGATGAAG 1380
Db 1441 CTACACCTCTGGCGGCTGGACCAAGGATGACATGTGCTCTAGACTGGCGGATGAAG 1500
QY 1381 CAGCAGCGGCTGTGCTGAGAGAGGGTCTCGATCCCAAGTCAACCATTTGTCATC 1440
Db 1501 CAGCAGCGGCTGTGCTGAGAGAGGGTCTCGATCCCAAGTCAACCATTTGTCATC 1560
QY 1441 TTTCCTGCTCCCATGTTATGCTGCCCCACAGAGTCCAGTGGCAGCTGCAGGTCCTCG 1500
Db 1561 TTTCCTGCTCCCATGTTATGCTGCCCCACAGAGTCCAGTGGCAGCTGCAGGTCCTCG 1620
QY 1501 ATGATTGCGGGTGCCCAATTTCTATGTTGGGAGGAGCCCTGCAGGAATGCCCATCCT 1560
Db 1621 ATGATTGCGGGTGCCCAATTTCTATGTTGGGAGGAGCCCTGCAGGAATGCCCATCCT 1680
QY 1561 GAAACCAAGAGATCTGATGAACCCACTATCGGGGCAAGTCTTGAGCATGGCCCT 1620
Db 1681 GAAACCAAGAGATCTGATGAACCCACTATCGGGGCAAGTCTTGAGCATGGCCCT 1740
QY 1621 GGCCCTCACCTCTGTGAAATCATTCATTCGAGTGGCTGCCCTACACAAAGCCAAAAA 1680
Db 1741 GGCCCTCACCTCTGTGAAATCATTCATTCGAGTGGCTGCCCTACACAAAGCCAAAAA 1800
QY 1681 GCCATGCACTCTPATGATCCAGCAAGGCAATGAGTTGACTTCATCTCAGGAATCGA 1740
Db 1801 GCCATGCACTCTPATGATCCAGCAAGGCAATGAGTTGACTTCATCTCAGGAATCGA 1860
QY 1741 ATGAGGAGCTGCCCGGAGGAGAGATTCCTCCAGATGGCTTCATGCCCCCAAGCA 1800
Db 1861 ATGAGGAGCTGCCCGGAGGAGAGATTCCTCCAGATGGCTTCATGCCCCCAAGCA 1920
QY 1801 TGGAGGCTGCTGACATTTATACAGTCCCTGGAGAAAGTA 1845
Db 1921 TGGAGGCTGCTGACATTTATACAGTCCCTGGAGAAAGTA 1965

RESULT 3
AF313907

LOCUS AF313907 1845 bp mRNA linear PRI 01-FEB-2001
DEFINITION Homo sapiens 3'-phosphoadenosine 5'-phosphosulfate synthase 2 alpha
mRNA, complete cds.
ACCESSION AF313907
VERSION AF313907.1 GI:12642583
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1845)
AUTHORS Fuda,H., Shimizu,C. and Strott,C.A.
TITLE Human Bifunctional 3'-Phosphoadenosine 5'-phosphosulfate Synthase:
Differential Expression of Isoforms and Effect of Polymorphisms on
Activity
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1845)
AUTHORS Shimizu,C., Fuda,H. and Strott,C.A.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2000) Section on Steroid Regulation, NICHD/NIH,
9000 Rockville Pike, Bethesda, MD 20892-4510, USA

FEATURES
source Location/Qualifiers
1..1845
/organism="Homo sapiens"
/db_xref="taxon:9606"

CDS
1..1845
/function="catalyzes production of 3'-phosphoadenosine
5'-phosphosulfate, the universal sulfonate donor molecule,
from inorganic sulfate and ATP"

/codon_start=1
/product="3'-phosphoadenosine 5'-phosphosulfate synthase 2
alpha"
/protein_id="AAK00296.1"
/db_xref="GI:12642584"

/translation="MSGIKQKTKENQOKSTNVVQAHVHVRNKRGVVGTGGPRGCT
VWLTGSGAGKTTISPALEYLYSHAIPCISLDGDVHRHGLNRLNGSPGDRENIR
IADVKLFADAGLVCTISFSPAKORENARKIHESAGLFFEIFYDAPLNCEISRDV
KGLYKRAAGEIKGTGIDSDYKPEPTPERVLTNLSTVSDCVHQVVVELLOENIVY
TIIKDHELFPENKLDHRAEATLPSLITLDLQWQVLSQVGVVVELLOENIVY
YLVQHFDTLLDDGVINMSPIVLVPSAEDTKLEGSKFLAHGRRVAILLRDAEY
EHKREKRCRWGVTCTTCKPHIKWVMSGDMVLGGDLQVLEKIRWPDGLQYRLTPLE
LKQCKEMNADAVFAQLRNPVHALLMQDTRRRLLRGYKHPVLLHLLPLGGWTKD
DVLPLWRMKAHAALVEGLDPKSTIVAIFPSPMLYAGTPEVQWHCRSRMIAGANFY
IVGRDPAGHPHETKLDIYPTGGKSVLSMAPGLTSVEIIPFRVAATNKAKKAMDFFD
PARHNEFDFTSGTRMKLAREGENPPDGMAPKAWKVLTYDYSLEKN"

BASE COUNT 494 a 436 c 496 g 419 t

Query Match 99.7%; Score 1840.2; DB 9; Length 1845;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1842; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATCTCGGGGATCAAGAGCAAAAGACGAGCAACAGCAAGAAATCCACCAATGTAGTCTAT 60
Db 1 ATCTCGGGGATCAAGAGCAAAAGACGAGCAACAGCAAGAAATCCACCAATGTAGTCTAT 60
QY 61 CAGGCCCAACCATCTGAGCAGGAATAAGAGAGGCAAGTGGTTGGAACAAGGGTGGGTTTC 120
Db 61 CAGGCCCAACCATCTGAGCAGGAATAAGAGAGGCAAGTGGTTGGAACAAGGGTGGGTTTC 120
QY 121 CGAGGATGTACCGTGGCTTAACAGGTCTCTCTGGTGTCTGGAACAAACAGTAAGTTTT 180
Db 121 CGAGGATGTACCGTGGCTTAACAGGTCTCTCTGGTGTCTGGAACAAACAGTAAGTTTT 180
QY 181 GCCTCTGAGGAGTACCTTGTCCTCCATGCCATCCCTTGTTACTCTCCCTGGATGGGACAAT 240
Db 181 GCCTCTGAGGAGTACCTTGTCCTCCATGCCATCCCTTGTTACTCTCCCTGGATGGGACAAT 240
QY 241 GTCCGTCATGGCCTTAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGGAAATATC 300
Db 241 GTCCGTCATGGCCTTAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGGAAATATC 300
QY 301 CGCCGATGTGCTGAGGTGGCTAAGCTGTGTGCTGATGCTGGTCTGCATTACCAGC 360

Db 1429 GACTGGCGGATGAGCAGCAGCGGCTGTGCTCGAGGAAGGGTCTCGGATCCCAAGTCA 1488

Qy 1426 ACCATTGTTGCCATCTTTCCTCCATCTTCCATCTTATATGCTGCGCCACAGAGGTCAGTGG 1485

Db 1489 ACCATTGTTGCCATCTTTCCTCCATCTTCCATCTTATATGCTGCGCCACAGAGGTCAGTGG 1548

Qy 1486 CACTGCAGGTCCTGGGATGATGCGGGTGCCAAATTTCTACATTTGCGGAGGGACCCCTGCA 1545

Db 1549 CACTGCAGGTCCTGGGATGATGCGGGTGCCAAATTTCTACATTTGCGGAGGGACCCCTGCA 1608

Qy 1546 GGAATGCCCATCTCTGAACCAAGAGGATCTGTATGAACCACTCATGCGGGGCAAGGTC 1605

Db 1609 GGAATGCCCATCTCTGAACCAAGAGGATCTGTATGAACCACTCATGCGGGGCAAGGTC 1668

Qy 1606 TTGAGCATGCCCTGCGCTCACCTCTGTGGAATCATTCATTCGAGTGGCTGCCTAC 1665

Db 1669 TTGAGCATGCCCTGCGCTCACCTCTGTGGAATCATTCATTCGAGTGGCTGCCTAC 1728

Qy 1666 AACAAAGCCAAAGAGGATGAGTGGATCTATGATCCAGCAGGACATGAGTTGACTTC 1725

Db 1729 AACAAAGCCAAAGAGGATGAGTGGATCTATGATCCAGCAGGACATGAGTTGACTTC 1788

Qy 1726 ATCTCAGGAATCGAATGAGGAGCTCGCCGGAAGGAGAGAAATCCCCAGATGGCTTC 1785

Db 1789 ATCTCAGGAATCGAATGAGGAGCTCGCCGGAAGGAGAGAAATCCCCAGATGGCTTC 1848

Qy 1786 ATGCCCCCAAGCATGGAAGGTCCTGACAGATTTATACAGGTCCTGGGAGAGAACTAA 1845

Db 1849 ATGCCCCCAAGCATGGAAGGTCCTGACAGATTTATACAGGTCCTGGGAGAGAACTAA 1908

RESULT 6

AF150754
LOCUS Homo sapiens 3'-phosphoadenosine 5'-phosphosulfate synthase 2b 2001
DEFINITION isoform mRNA, complete cds.
ACCESSION AF150754
VERSION AF150754.2 GI:12484558.

SOURCE

ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Venkatchalam, K.V., Fuda, H. and Strott, C.A.
TITLE 3'-phosphoadenosine 5'-phosphosulfate synthase 2b isoform
JOURNAL Unpublished

REFERENCE

AUTHORS Venkatchalam, K.V., Fuda, H. and Strott, C.A.
TITLE Direct Submission
JOURNAL Submitted (12-MAY-1999) Steroid Regulation, National Institute of Health, 9000 Rockville Pike, Bethesda, MD 20892-4510, USA

REFERENCE

AUTHORS Venkatchalam, K.V., Fuda, H. and Strott, C.A.
TITLE Direct Submission
JOURNAL Submitted (25-JAN-2001) Steroid Regulation, National Institute of Health, 9000 Rockville Pike, Bethesda, MD 20892-4510, USA

REMARK

COMMENT Sequence update by submitter
On Jan 25, 2001 this sequence version replaced gi:6642928.

FEATURES

source Location/Qualifiers
1..2377
/organism="Homo sapiens"
/db_xref="taxon:9606"

CDS

81..1940
/note="PAPS synthase 2b b; ATP sulfurylase/adenosine 5'-phosphosulfate kinase 2b; bifunctional enzyme that regulates production of the universal sulfonate donor molecule, 3'-phosphoadenosine 5'-phosphosulfate"
/codon_start=1
/product="3'-phosphoadenosine 5'-phosphosulfate synthase 2b isoform"
/protein_id="AAF20366.2"
/db_xref="GI:12484559"

/translation="MSGIKKKQKTNQOKSTNVVQAHVSRNKRKQVVGTRGGRGCT
WLTGSLGAGKTTISFALEYLVSHAIPCYSLDGNVNRHGLNRLNGSPGDRENIIR
IAEVAKLFADAGLVITFSIFPAKDRNKAITHESAGLFEFFEDVADINLCSRRD
KGLYKRRAGEIKGTGIDSDYKPTPEPVRLKTNLSTSDCVHVOVELLEQONIPY
TIIKDHELEVPENKLDHVAEATLPSLSTTKLDLQWVOVLSGFWATPLKGFMRKE
YLOVMFDTLDDGHALPDGVINMSIPIVPSAEDKTRLEGCKSFVLAHGRRVAILR
DAFEYHRKEERSRVMTGTTCTKHPHKNHWSGDWLVDGLQVLEKIRWNGDLQYR
LTFLELKQCKEMNADAVAFQLRPNVNGHALMDQTRRLRLGTYKHVLLHLLPLG
GWTKDDVPLDWRMKHAAVLEGVLPKSTIVAFPSPLMAYAGTVPQVHCRSMIA
GANFYLVGRDPAGMPHETPKKDLVEPTHTGKVLSPGLTSVEIIPRVAAYNRKAKA
MDFYDPAHNEFDIFISGTRMRKLAREGENPDGFMAPKANVKLVLDIYRSLEKN"

polyA_site 636 a 556 c 584 g 601 t
BASE COUNT
ORIGIN

Query Match 98.3%; Score 1813.6; DB 9; Length 2377;
Best Local Similarity 99.08; Pred. No. 0;
Matches 1841; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

Qy 1 ATGTGCGGGATCAAGAAAGCAAGACCGAGAACCCAGAGAAATCCACCAATGTAGTCTAT 60

Db 81 ATGTGCGGGATCAAGAAAGCAAGACCGAGAACCCAGAGAAATCCACCAATGTAGTCTAT 140

Qy 61 CAGGCCACCATGTGACAGAGAAATAGAGAGGCAAGTGGTGAACAAGGGTGGCTTC 120

Db 141 CAGGCCACCATGTGACAGAGAAATAGAGAGGCAAGTGGTGGAAACAGGGTGGCTTC 200

Qy 121 CGAGGATGTACCGTGTGGCTAACAGGTCCTCTGTGGTCTGGAACCAACGATAAGTTT 180

Db 201 CGAGGATGTACCGTGTGGCTAACAGGTCCTCTGTGGTCTGGAACCAACGATAAGTTT 260

Qy 181 GCCTCGAGGAGTACCTTGTCTCCATGCCATCCCTTGTACTCCCTGGATGGGACAAT 240

Db 261 GCCTCGAGGAGTACCTTGTCTCCATGCCATCCCTTGTACTCCCTGGATGGGACAAT 320

Qy 241 GTCCGTCATGCCCTTAACAGAAATCTCGGATCTCTCTCTGGGACAGAGAGAAATATC 300

Db 321 GTCCGTCATGCCCTTAACAGAAATCTCGGATCTCTCTCTGGGACAGAGAGAAATATC 380

Qy 301 CGCCGGATTGCTGAGGTGGCTAAGCTGTTTGTGCTGATGCTGCTGCTGCTGCATTACCAGC 360

Db 381 CGCCGGATTGCTGAGGTGGCTAAGCTGTTTGTGCTGATGCTGCTGCTGCATTACCAGC 440

Qy 361 TTTATTTCCTCATTCGCAAGGATCTGAGAAATGCCGCAAAATACATGATCAGCAGG 420

Db 441 TTTATTTCCTCATTCGCAAGGATCTGAGAAATGCCGCAAAATACATGATCAGCAGG 500

Qy 421 CTGCCATTCTTCGAAATATTTGTAGATGCACCTCTTAATATTTGTGAAGCAGAGACGTA 480

Db 501 CTGCCATTCTTCGAAATATTTGTAGATGCACCTCTTAATATTTGTGAAGCAGAGACGTA 560

Qy 481 AAAGGGCTCTATAAAAGGGCCAGAGCTGGGAGATTAAGGATTTACAGGATTTGATTCT 540

Db 561 AAAGGGCTCTATAAAAGGGCCAGAGCTGGGAGATTAAGGATTTACAGGATTTGATTCT 620

Qy 541 GATTATGAGAAACCTGAAACTCTCTGACCGTGTCTTAAACCAATTTGTCCACAGTCACT 600

Db 621 GATTATGAGAAACCTGAAACTCTCTGACCGTGTCTTAAACCAATTTGTCCACAGTCACT 680

Qy 601 GACTGTGTCACCCAGGTAGTGAACCTCTTCAAGAGCAGAACATTTGTACCTTACTATA 660

Db 681 GACTGTGTCACCCAGGTAGTGAACCTCTTCAAGAGCAGAACATTTGTACCTTACTATA 740

Qy 661 ATCAAGATATCCACGAACCTTTTGTGCCGGAACCAACATTTGACCACTCCAGGCTGAG 720

Db 741 ATCAAGATATCCACGAACCTTTTGTGCCGGAACCAACATTTGACCACTCCAGGCTGAG 800

Qy 721 GCTGAACCTCTCCCTTCATTATCAATTAAGTAACTCCAGTCCAGTGGGTCAGGTTTG 780

Db 801 GCTGAACCTCTCCCTTCATTATCAATTAAGTAACTCCAGTCCAGTGGGTCAGGTTTG 860

Qy 781 AGCGAAGGCTGGGCCACTCCCTCAAGAGGTTTTCATCGGGAGAGAGGAGTACTTACAGTT 840


```

552 CTACAGAGAGCGCGGAGGAGATATAAGGCTTCACTGGCATCGATTCTCAGTATGA 611
549 GAAACCTGAAACTCCTGAGGCTGTCTTAAACCAATTTGTCACACAGTGCACGTGTGT 608
612 GAAACCGAGGCGCGGAGTGTGTGTAACCGGATTTCTGTGACGTCAACAGTGCCT 671
609 CCACAGGTAGTGGAACTTCTGCAAGAGCAGACATTTGTACCCCTATATAATCAAGA 668
672 CCAGCAGGTTGTGAGCTTCTTCCAGGAACGGGACATCGTCCCTGTGGATGCTTCCATGA 731
669 TATCCAGCACTCTTTGTCGCGGAAACAAACATTTGACACAGTCCGAGCTGAGCTGAAC 728
732 AGTGAAGAGCTATATGTCCAGAGATAAATCTTCTGGCCAAACACTGATCAGAACG 791
729 TCTCCCTTCAATTAATCAATTAAGCTGGATCTCCAGTGGTCCAGGTTTGGAGCAAGG 788
792 CTTACACAGCCTGAAATCAATTAAGTGGATATGCAAGTGGTGGTGGTGGTGGCAAGG 851
789 CTGGCCACTCCCTCAAGGTTTTCATGCGGGAGAGAGTACTTACAGGTTATGCACTT 848
852 TTGGCGGACTCCTCTGAACGGCTTCATGAGAGAGGGAGTACTTGCAGTGCCTCCATTT 911
849 TGACACCCCTAGATGATGGCTGTATCAACATGACATCCCATTTGTACTGCCCTCTC 908
912 CGATTGTCTTGGATGGAGGCGTCACTCACTTATCGGTGCTTATAGTTCTGACAGCTAC 971
909 TGCAGAGGTAAGACACGCTGGAAGGTCACAGAGTTTGTCTGGCACATGGTGGACG 968
972 GCACGAGGATAAGAGAGGCTGCACGGCTGCACCGGTTCTGCTGTGTATGAGGGCG 1031
969 GAGGAGTATCTTACGAGACCTGAATCTATGAACACAGAAAGAGACGCTGTTC 1028
1032 CCCTGGGCGCATCTTCGGAATCCTTGAATTTTGGAGCAGCGGAGAGGCGGTGTC 1091
1029 CCGTGTGTTGGGACACATGTACAAACACCCCATATCAAAATGATGATGAAAGTG 1088
1092 CAGAGTGGGAAACACATGCAAGAACCCCTCATCAAGATGGTCTGGAACAGG 1151
1089 GGACTGGCTGTTGGAGACCTTCAAGTGTCTGGAGAAATAGATGAATGATGGCT 1148
1152 GGATTGGCTGATGGAGGAGATCTCAAGTCTCTGGACCGATTTACTGGAATGATGCT 1211
1149 GGACCAATACCTCTGACACCTCTGGAGCTCAACAGAAATGTAAAGAAATGAATGCTGA 1208
1212 TGATGATACCGCTTTACCCCGACGGAGCTCAAGCAGAAAGTTTAAAGATATGAACGCTGA 1271
1209 TGGGTGTTTGGATTCCAGTTGGCAATCTGTGCCAATGCTGCAATGCTGCTGTGATGCA 1268
1272 TGTGTCTTTGCAATTTGAGTTGGCAACCCAGTGCACACGGGCGCTCTGTTAATGCA 1331
1269 GGACACTCTGCGCAGGCTCTAGAGAGGGCTACAGACACCCGCTCTCTACTACACCC 1328
1332 GGATACCCACACGCTTCTGGAGAGGGCTACCGGCGCTGCTCTCTCTCCATCC 1391
1329 TCTGGCGGCTGACCAAGATGACGATGCTCTAGACTGGCGGATGAAGCAGCACGC 1388
1392 TCTTGTGGCTGACCAAGGATGACGATGCTCTGATGTCGCTGATGAGCAGCACGC 1451
1389 GCGTGTCTGAGGAGGGTCTCTGATCCCAAGTCAACCATGTTGCCATCTTCCGTC 1448
1452 TGCAGTGTGGAGAGGGATCTCTGATCTTGAACGACAGTGGTGGCCATTTTCCCGTC 1511
1449 TCCCATGTTATATGCTGGGCCACAGAGGTCCAGTGGCACATGCAAGTCCCGGATGTC 1508
1512 TCCTATGATGATGCTGGCCAAACCGAGGTCCAGTGGCACTGCAGGCGCGGATGTCG 1571
1509 GGTGCCAATTTCTACATGTTGGGAGGAGCCCTGCAGAAATGCCCTCCCTGAACACAA 1568
1572 CGAGGCCAAATTTTATATTGTTGGAGAGACCCCTGCTGGCATGCTCCCGGAGAGG 1631
1569 GAAGGATCTGTATGAACCACTCATGGGCGAGGTCCTTGAGCATGCGCCCTGGCCTCAC 1628
1632 GAAGGACCTCTATGAGCCAAACACATGTTGCCAAAGTCTGACGATGCCCCAGGCTGAT 1691
1629 CTCTGTGGAATCATTCCTAGTGGCTGCTCAACAAAGCCAAAGAACCCATGGA 1688
1692 TACCTTGGAAATCGTTCCCTTCCGAGTTGCGAGTTTACACAAAGAAAGAGCGGATGGA 1751
1689 CTCTTATGATCCAGCAAGCACAATGAGTTTGACTTCTATCTCAGCAACTCGAATGAGAA 1748
1752 CTACTATGACTCTGAGCACCACGAACTTCGAGTTTATTTTTCAGGAACGAGGATGCGCAA 1811
1749 GCTCGCCCGGAGGAGAGAAATCCCGAGATGCTTCATGTCGCCCCCAAGCACTGGAAGT 1808
1812 GCTGGCAGCGAAGGCGCAGAACTCCAGAGGCTTCATGCCCCCAAGGCTGGAAGT 1871
1809 CCGTGCAGATATTATACAGGTCCTCGGAGAGAACTAA 1845
1872 GCTGGTAGACTACTACAAGTCTTAGAAGAACCTAA 1908

RESULT 11
LOCUS E27970 1875 bp DNA linear PAT 18-JUN-2001
DEFINITION APS kinase/ATP sulfurylase gene derived from human being.
ACCESSION E27970.1 GI:13026556
VERSION JP 1999187883-A/1.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1875)
AUTHORS Katsuhiko, T., Hiroshi, N. and Masahito, S.
TITLE APS kinase/ATP sulfurylase gene derived from human being
JOURNAL Patent: JP 1999187883-A 1 13-JUL-1999;
COMMENT JAPAN HEALTH SCIENCES FOUNDATION, UNITIKA LTD
PN JP 1999187883-A/1
PD 13-JUL-1999
PF 26-DEC-1997 JP 1997360387
PR KATSUHIKO TAKAYANAGI, HIROSHI NAKAJIMA, MASAHIRO SUIKO PC
PI C12N15/09, C12N1/21, C12N9/12, C12P17/18, C12N15/09, C12R1:91, PC
(C12N1/21, C12R1:19), (C12N9/12, C12R1:19), (C12N15/00, C12N15/00, PC
C12R1:91)
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..1875
FT source Location/Qualifiers
FT source 1..1875
FT source /organism="Homo sapiens"
BASE COUNT 526 a 379 c 499 g 471 t
ORIGIN
Query Match 53.7%; Score 990; DB 6; Length 1875;
Best Local Similarity 71.3%; Pred. No. 8.6e-272;
Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;
QY 9 GATCAGAGCAAAACACGAGAGACCGAGAACCCAGCAGAAATCCACCATGTAGTCTATCAGGCCCA 68
DB 39 GAGCAATAACGCGCAACTGGGAAATGCAGAGAGCAACCAATGTCCCTACCAAGCCCA 98
QY 69 CATCTGAGCAGGAATAAGAGAGGCGCAAGTGGTTGGAACAAGGGGTGGGTCCGAGGATG 128
DB 99 TCATGTACAGGGAACAAGAGAGTCAAGTGGTGGGACCAAGAGGTGGCTTCTGCTGTTG 158
QY 129 TACCGTGTGCTAACAGGTCTCTCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 188
DB 159 CACAGTTGGCTAACAGGCTTGTCTGAGCGGGAAGAGACTACTGTGAGCATGCGCTTGA 218
QY 189 GGAGTACCTTGTCTCCCATGCCATCCCTCTGTTACTCCCTGGATGGGACAAATGTCCTGCA 248
```

Db	1299	GGATACCCATAAGCAACTTCTAGAGAGAGGGGCTACCGGCGCCCTGTCTCTCTCTCCACACC	1355
QY	1329	TCTGGCGGCTGGACCAAGGATGACGATGTGCCTTCTAGACTGGCGGATGAAGCAGCAGC	1388
Db	1359	TCTGGGTGCTGGACAAGGATGAGGATGTTCCCTTGTATGGCGTATGAAGCAGCATGC	1418
QY	1389	GGCTGTGCTCGAAGAGGGTCTCGGATCCCAAGTCAACCAATGTTGCCAFCTTTCCGTC	1448
Db	1419	TGCAGTGTGGAGGAAGAGTCTGTAATCTCTGAGACGACAGTGGTGGCCATCTTCCCATC	1478
QY	1449	TCCCATGTTATCTCTGGCCCCACAGAGGTCCTGAGTGGCACTGCGAGTCCCGGATGATGC	1508
Db	1479	TCCCATGATGATCTGTGACCAAGTCCAGTGGCATTTGCGAGACCGGATGGTGTGC	1538
QY	1509	GGGTGCCAATTTCTACATTTGGGGAGGGACCTTCAGGAATGCCCATCTCTGAAACCAA	1568
Db	1539	AGGAGCCCACTTTTACATTTGTGACGAGACCTCTGCTGGCATGCTCATCCAGAAACAGG	1598
QY	1569	GAAGGATCTGTGAACCCACTCATGTTGGGGCAAGTCTTGAGCATGCGCCCTGCGCTCAC	1628
Db	1599	GAAGGATCTTTATGAGCCAAGTCATGTTGCGCCAAAGTGTGACGATGGCCCTGGTTAAT	1658
QY	1629	CTCTGTGGAATCATTTCCATTCGAGTGGCTGCCTACACAAAGCCAAAAGCCATGGA	1688
Db	1659	CACTTTGGAAATAGTTCCCTTTCAGTTGCACTTACAAAGAAAAGAGCGTATGGA	1718
QY	1689	CTTCTATGATCCAGCAAGGCAACATGAGTTTGTACTTTCATCTCAGGAACCTCGAATGAGAA	1748
Db	1719	CTACTATGACTCTGAACACCATGAAGACTTTGAATTTATTTTCAGGACAGGAATGCGCAA	1778
QY	1749	GCTCGCCCGGGAAGAGAGAAATCCCCAGATGGCTTCATGCGCCCCCAAGCATGGAAGGT	1808
Db	1779	ACTTGTCTGAGAAGGCCAGAACACCTGAAGGTTTCATGGCTCCCAAGGCTTGGACCGT	1838
QY	1809	CTGACAGATTTATTACAGGTCCTCTGGAGAA	1838
Db	1839	GCTGACAGAACTACTAAATCCTTGGAGAA	1868
RESULT 12			
AF033026		2265 bp	linear
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			

27. .1901
/dev -slay

Db 1840 CACTTTGGAATAGTTCCTTTTCAGTTGCGAGCTTACACAAAGAAAAGAGCGCATGGA 1899
Qy 1689 CTTCTATGATCCAGCAGGACCAATGAGTTTGAATCTTATCTCAGGAATCGAATGAGGAA 1748
Db 1900 CTACTATGACTCTGAACACCATGAGACTTTGAATTTATTCAGACACCAATGCGGAA 1959
Qy 1749 GCTCGCCCGGGAAGAGAGATCCCCAGATGGCTTCATGCCCCCAGGAGCATGGAAGGT 1808
Db 1960 ACTTGCTCGAAGGCGGAGAACCAACCATGGAAGTTTCATGGCTCCCAAGGCTTGACCGT 2019
Qy 1809 CTTGACAGATATTACAGGTCCTCGGAGAA 1838
Db 2020 GTCGACAGATATTACAAATCCTTGGAGAA 2049

RESULT 15
BC011392
LOCUS BC011392 2430 bp mRNA linear PRI 30-JUL-2001
DEFINITION Homo sapiens, Similar to PAPS synthetase, clone MGC:9898
IMAGE:3869484, mRNA, complete cds.
ACCESSION BC011392
VERSION BC011392.1 GI:15030251
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2430)
Strausberg, R.
Direct Submission
Submitted (25-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590;
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalobebcm.tmc.edu
Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC plate: 13 Row: 1 Column: 18
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES
source Location/Qualifiers
1. .2430
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:9898 IMAGE:3869484"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH-MGC-67"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
21. .1832
/codon_start=1
/product="Similar to PAPS synthetase"
/protein_id="AAH11392.1"
/db_xref="GI:15030251"
/translation="MQRATNVTYQAHVSRNKGQVVGTRGGRGCTVWLTGLSGACK
TTVSMALKEVLCVHIPCYTLGDNIQGLNKLGFSPDEENVRRIAEVAKLFADA
GLVCITFSIPYDORNRNAQIHGASLIPFEVFDAPLVHCEORDVGLKFKKARAGE
IKGTGTIDSEYEKPEAPLVLKTDSCDNDVCVQVVVLLQERDIPVDASVEVKELV

PENKHLIAKTAETLPAKINKVDMQWQVLAEGWATPLNGFREREYLOLCHDFCLL
DGGVNLISVPIVLTATHEKDLGCTAFALMYERRVAILRNEFFEPHREKRCARQ
WGTTCKNHPYIKVMEOCDWLGDLQVLDVYNDGLOVRLAFTKOKFKDMNAD
AVFAFOLRNPNVHNGHALLMODTHKQLERGRYRPPVLLHLLPLGGTCKDDVPLMRMKQ
HAALVEGVLNPETTVVAIFSPMMIAGTEVQWCHRMKRVAGANFYILVGRDPAGMPH
PETKDKLYEFSHGAKVLTMLPLITLEIVPFRVAAYNRKKKKRMDYDSEHEDFEFTS
GTRMRKILAREGQKPEGFMAKWTVEITYYKSLKA"

BASE COUNT 688 a 475 c 567 g 700 t
ORIGIN

Query Match 53.6%; Score 988.6; DB 9; Length 2430;
Best Local Similarity 71.8%; Pred. No. 2.2e-271;
Matches 1294; Conservative 0; Mismatches 501; Indels 0; Gaps 0;
Qy 36 GCAGAAATCCACCAATGTAGTCTATCAGGCGCCACCATGTGACGACGAGTAAGAGAGGCA 95
Db 23 GCAGAGACCAACCAATGTACCTACCAAGCCCATCATGTGACGAGCAACAGAGAGTCA 82
Qy 96 AGTGGTTGGAACAAGGGTGGTTCGAGGATGTACCGTGTGGCTTAACAGGTCCTCTGG 155
Db 83 GGTGGTGGGACACAGAGGTGGCTTCGTGGTTGCACAGTTTGGCTTAACAGGCTTCTCTGG 142
Qy 156 TGCTGGAACCAACAGGATAAGTTTGGCCCTGGAGGAGTACCTTGTCTCCCATGCCATCCC 215
Db 143 AGCGGGAAGAGTACTGTGAGCATGGCCCTGGAGGAGTACCTGGTGTTCATGGTATTCC 202
Qy 216 TTGTTACTCCCTGGATGGGACAAATGTCCGTCATGCCCTTAACAGAAATCTCGGATCTC 275
Db 203 ATGCTACACTCTGGATGGTGAACAATATTCGTCAGGCTCTCAATAAAATCTTGGCTTAG 262
Qy 276 TCCTGGGACAGAGAGAGAAATATCCGCCGATGTCTGAGGTGGCTTAAGCTTGTGGTGA 335
Db 263 TCCTGAAGACAGAGAGAGAAATGTTCGACGCGATCGCAGAAGTTGTAACATGTTTGCAGA 322
Qy 336 TGCTGGTCTGCTGCTACCTACCAGCTTTATTTCTCCATTCGCAAGGATCGTGAAGTGC 395
Db 323 TGCTGGCTTAGTGTGCATCACAAGTTTCATATCACCTTACATCAGGATCGCAACATGC 382
Qy 396 CCGCAAAATACATGAATCAGCAGGCTGCCATTTCTTTGAAATATTTGTAGATGCACTCT 455
Db 383 AAGCAAAATCATGAAGTGCAGTTTACCGTTTTCAGGATTTTGTGATGCTCTCT 442
Qy 456 AAATATTGTGAAAGCAGAGAGAGTAAAGGCTCTATAAAAGGCGCAGAGTGGGAGAT 515
Db 443 GCATGTTGTGAACAGAGGAGTGTCAAAGGACTCTACAAAAAGCCCGGAGGAGAAAT 502
Qy 516 TAAGGATTTACAGGTATTGATTCGTATTATGAGAACTGAACTCCCTGAGCGTGTCT 575
Db 503 TAAAGGTTTCACTGGGATCGATCTGAATATGAAAGCCAGAGGCCCCCTGATGGTGTCT 562
Qy 576 TAAACCAATTTGCCACAGTGAAGTACTGTGTCCACCAAGGTAGTGGAACTTTCTCAAGA 635
Db 563 GAAACAGACTCTGTGATGTAAATGACTGTGTCCAGCAAGTTGTGGAACTTCTACAGA 622
Qy 636 GCAGAACTTACCTTACTATATAATCAAGATATATCCAGAACTCTTTTGGCCGGAAGA 695
Db 623 ACGGATATTGTACCTGTGGATGATCTTATGAAGTAAAGAACTATATATGCCCCAGAAA 682
Qy 696 CAACCTTGACCACTCCGAGCTGAGGCTGAAGTCTCCCTTCATTTATCAATTAAGT 755
Db 683 TAACTTTCATTTGGCAAAACAGATGCGGAAACATTTACCACTGAAATTAATAGT 742
Qy 756 GGATCTCCAGTGGGTCCAGGTTTTCAGGCAAGGCTGGGCCACTCCCTCAAGAGTTTCAT 815
Db 743 GGATATGAGTGGGTGGAGGTTTGGCAGAGGTTGGCAACCCCATTAAGTGGCTTAT 802
Qy 816 GCGGAGAGAGGAGTACTTTACAGGTATATGCATTTTGACACCTCTGTAGATGATGGGTAT 875
Db 803 GAGAGAGAGGAGTACTTTCAGTGCCTTCATTTTGAATTTGTCTCTGATGGAGGTGTCAT 862
Qy 876 CAACATGAGCATCCCCATTTACTGCCGCTCTCTGACAGAGGATAGACACCGCTGGAAG 935
Db 863 TAACCTGTGACTACCTATAGTTCTGACTGCGACTCATGAAGATAAAGAGAGGCTGGACGG 922

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2003, 14:57:46 ; Search time 445 Seconds
(without alignments)
9336.937 Million cell updates/sec

Title: US-09-898-165B-9

Perfect score: 1845

Sequence: 1 atgtcgggagatcaagaagca.....ggtccttggaagaactaa 1845

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1813.6	98.3	2617	21	AA294206 Human transferase
2	990	53.7	1875	20	AA84897 Human APS Kinase/A
3	990	53.7	2507	19	AAV33482 Human disease rela
4	990	53.7	2527	20	AAV00717 Human secreted pro
5	990	53.7	2527	20	AAV00686 Human secreted pro
6	990	53.7	2646	24	AAV04855 Human DNA sequence
7	634	34.4	2463	23	ABU29667 Drosophila melanog
c	520.8	28.2	17196	23	ABU29666 Drosophila melanog
8	484.8	26.3	990	16	AAT04849 Human adenosine 5'

10	384	20.8	1695	21	AAC34705 Arabidopsis thalia
11	383	20.8	3684	21	AAC34705 SAMS promoter::ATP
12	378.8	20.5	401	21	AAH31044 Human colon cancer
13	378.8	20.5	401	21	AAH31145 Human colon cancer
14	360.8	19.6	1687	21	AAC47786 Arabidopsis thalia
15	359.8	19.5	1533	21	AAC42592 Arabidopsis thalia
16	357.6	19.4	1534	21	AAC47649 Arabidopsis thalia
17	351.2	19.0	1617	21	AAC39151 Arabidopsis thalia
c	315	17.1	574	24	ABN93862 Gene #360 used to
18	266.4	14.4	600	16	AAT04850 Human adenosine 5'
19	266.4	14.4	600	16	AAT04850 Human breast cance
20	233.4	12.7	435	22	AAL23324 Human breast cance
21	233.4	12.7	449	22	AAL14348 Human breast cance
c	233.4	12.7	490	22	AAL23210 Human breast cance
23	233.4	12.7	512	22	AAL14462 Human breast cance
24	233.4	12.7	517	22	AAL23345 Human breast cance
25	233.4	12.7	519	22	AAL14483 Human breast cance
c	26	23.4	13418	22	AAK86473 Human immune/haema
27	226	12.2	2323	19	AAV59524 Human secreted pro
28	206.4	11.2	692	21	AAC38756 Arabidopsis thalia
29	201.6	10.9	399	24	ABQ54857 Human ovarian anti
30	195.6	10.6	471	22	ABA59209 Human foetal liver
31	195.6	10.6	471	22	ABA27965 Probe #6431 for ge
32	195.6	10.6	471	22	AAK07411 Human brain expres
33	195.6	10.6	471	22	AAK33189 Human bone marrow
34	195.6	10.6	471	22	AAI38986 Probe #7672 used t
35	195.6	10.6	471	24	ABS08017 Human genome-deriv
36	195	10.6	195	22	ABA71730 Human foetal liver
37	195	10.6	195	22	ABA37826 Probe #16292 for g
38	195	10.6	195	22	AAK20089 Human brain expres
39	195	10.6	195	22	AAK46143 Human bone marrow
40	195	10.6	195	22	AAI52052 Probe #20738 used
41	195	10.6	195	24	ABS20450 Human genome-deriv
42	169.2	9.2	1160	21	AAA08597 Saccharomyces carl
43	165	8.9	222	21	AAA08749 Rat hepatocyte car
44	157.2	8.5	936	21	AAZ50162 Soybean Adenylisu
45	152	8.2	1217	21	AAZ50160 Corn Adenylisulph

ALIGNMENTS

RESULT 1

AAZ94206

ID AAZ94206 standard; cDNA; 2617 BP.

XX AC AAZ94206;

XX DT 19-JUN-2000 (first entry)

XX DE Human transferase TRNSFS-6 CDNA clone 1420940CB1.

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

XX DE Human transferase TRNSFS-6; human; antitumour; cancer; inflammation;

(INCY-) INCYTE PHARM INC.

Tang YT, Corley NC, Guegler KJ, Baughm MR, Lal P, Yue H;
Hillman JL, Azimzai Y;

WPI: 2000-256996/22.

P-PSDB; AAY79214.

Human transferase proteins useful for preventing, diagnosing and
treating cancers and developmental, gastrointestinal, genetic,
immunological, neurological, reproductive and smooth muscle disorders -
Claim 9; Page 100-101; 113pp; English.

The present sequence is that of cDNA clone 1420940C81 encoding
human transferase TRNFS-6 (see AAY79214), 1 of 15 claimed human
transferase proteins of the invention (see AAY79209-23). The clone
was isolated from foetal kidney cDNA library KIDNN009. TRNFS-6
is expressed in cardiovascular, gastrointestinal, reproductive and
endocrine tissues, especially those associated with cancer,
inflammation and cell proliferation. It shows homology to ATP
sulfonylase/APS kinase 2. A polynucleotide comprising nucleotides
942-986 of the present sequence can be used as a DNA probe. The
new human transferase proteins and the polynucleotides encoding
them can be used in the diagnosis, prevention and treatment
(including gene therapy and antisense therapy) of cancer,
developmental disorders, gastrointestinal disorders, genetic
disorders, immunological disorders, neurological disorders,
reproductive disorders, and smooth muscle disorders.

SQ Sequence 2617 BP; 709 A; 602 C; 630 G; 676 T; 0 other;

Query Match 98.3%; Score 1813.6; DB 21; Length 2617;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 1841; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

PA	633	GATTATGAGAAACCTGAAACCTCTGAGCGTGTGCTTAAACCAATTTGTCCACAGTGAGT	692
XX			
PI	601	GACTGTGTCACACAGGTAGTGAACCTTCTGCAAGAGCAGAACATTTGTACCTTATACATATA	660
XX			
DR	693	GACTGTGTCACACAGGTAGTGAACCTTCTGCAAGAGCAGAACATTTGTACCTTATACATATA	752
XX			
PT	661	ATCAAAAGATATCACGAACCTCTTTGTGCGCGGAAACAAACTTGACCACGCTCCAGAGTGAG	720
XX			
PT	753	ATCAAAAGATATCACGAACCTCTTTGTGCGCGGAAACAAACTTGACCACGCTCCAGAGTGAG	812
XX			
PS	721	GCTGAAACCTCTCCCTTCATTTATCAATTAAGTAACTGATCTCCAGTGGGTCCAGGTTTG	780
XX			
CC	813	GCTGAAACCTCTCCCTTCATTTATCAATTAAGTAACTGATCTCCAGTGGGTCCAGGTTTG	872
XX			
CC	781	AGCGAAGGCTGGGCCACTCCCTTCAAGGTTTCATGCGGAGAGAGGAGTACTTTACAGGTT	840
XX			
CC	873	AGCGAAGGCTGGGCCACTCCCTTCAAGGTTTCATGCGGAGAGAGAGTACTTTACAGGTT	932
XX			
CC	841	ATGCACTTTGACACCTGCTAGAGTCTCTGAGAGGAGTCTTACGAGACGCTGAATTTCTATGAA	885
XX			
CC	933	ATGCACTTTGACACCTGCTAGAGTCTCTGAGAGGAGTCTTACGAGACGCTGAATTTCTATGAA	992
XX			
CC	886	ATCCCATTTGACTGCGCGTCTCTGAGAGGATTAAGACACGCGTGGNAGGTCGACGAAG	945
XX			
CC	993	ATCCCATTTGACTGCGCGTCTCTGAGAGGATTAAGACACGCGTGGNAGGTCGACGAAG	1052
XX			
CC	946	TTTGTCTCTGGCACATGTTGACGAGGAGTCTTACGAGACGCTGAATTTCTATGAA	1005
XX			
CC	1053	TTTGTCTCTGGCACATGTTGACGAGGAGTCTTACGAGACGCTGAATTTCTATGAA	1112
XX			
CC	1006	CACAGAAAAGAGAACGCTGTTCCCGTGTGTTGGGGACACATGTACAAACACCCCAT	1065
XX			
CC	1113	CACAGAAAAGAGAACGCTGTTCCCGTGTGTTGGGGACACATGTACAAACACCCCAT	1172
XX			
CC	1066	ATCAAAATGGTGAATGAAAGTGGGACTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAG	1125
XX			
CC	1173	ATCAAAATGGTGAATGAAAGTGGGACTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAG	1232
XX			
CC	1126	AAAAAAGATGGAATGATGGGTGGACCAATACCGTCTGACACCTCTGGAGCTCAACACAG	1185
XX			
CC	1233	AAAAAAGATGGAATGATGGGTGGACCAATACCGTCTGACACCTCTGGAGCTCAACACAG	1292
XX			
CC	1186	AAATGTAAGAATGAATGCTGATGCGGTGTTTGGCAATCCAGTTGGCAATCTCTGCCAC	1245
XX			
CC	1293	AAATGTAAGAATGAATGCTGATGCGGTGTTTGGCAATCCAGTTGGCAATCTCTGCCAC	1352
XX			
CC	1246	AATGGCCATGCCCTGTTGATGAGGACACCTGCCGAGGCTCTAGAGAGGGCTACAAG	1305
XX			
CC	1353	AATGGCCATGCCCTGTTGATGAGGACACCTGCCGAGGCTCTAGAGAGGGCTACAAG	1412
XX			
CC	1306	CACCCGGTCTCTACTACACCTCTGGGCGGTGGACCAAGGATGACGATGTCCTCTA	1365
XX			
CC	1413	CACCCGGTCTCTACTACACCTCTGGGCGGTGGACCAAGGATGACGATGTCCTCTA	1472
XX			
CC	1366	GACTGCGGATGAAGCAGACGCGGTGCTGCGAGAGAGGGTCTCTGGATCCCAAGTCA	1425
XX			
CC	1473	GACTGCGGATGAAGCAGACGCGGTGCTGCGAGAGAGGGTCTCTGGATCCCAAGTCA	1532
XX			
CC	1426	ACCATGTTGCCATCTTCCGCTCCCATGTTATATGCTGGCCCCACAGAGTCCAGTGG	1485
XX			
CC	1533	ACCATGTTGCCATCTTCCGCTCCCATGTTATATGCTGGCCCCACAGAGTCCAGTGG	1592
XX			
CC	1486	CACCTGAGTCCCGGATGATGGGGTGCCAAATTTCTACATTTGTGGGGAGGACCCCTGCA	1545
XX			
CC	1593	CACCTGAGTCCCGGATGATGGGGTGCCAAATTTCTACATTTGTGGGGAGGACCCCTGCA	1652
XX			
CC	1546	GAATGCCCCATCTCGAAACCAAGAGGATCTGTATGAACCCACTCATCTGGGGCAAGTTC	1605
XX			
CC	1653	GAATGCCCCATCTCGAAACCAAGAGGATCTGTATGAACCCACTCATCTGGGGCAAGTTC	1712
XX			
CC	1606	TTGAGCATGGCCCTGGCCTACCTCTGTGGAATCATTTCCATTCGAGTGGCTGCTAC	1665
XX			

Db	1713	TTGAGCATGGCCCTGGCCCTACCTCTGTGGAAATCATTCATTCAGTGGCTGCCTAC	1777
Qy	1666	ACAAAAGCCAAAAGCCATGGACTTCTATGATCCAGCAAGGCACAAATGAGTTTGACTTC	1725
Db	1773	ACAAAGCCAAAAGCCATGGACTTCTATGATCCAGCAAGGCACAAATGAGTTTGACTTC	1832
Qy	1726	ATCTCAGGAACTCGAATGAGGAAGCTCGCCCGGAAGGAGAGAATCCCCAGATGGCTTC	1785
Db	1833	ATCTCAGGAACTCGAATGAGGAAGCTCGCCCGGAAGGAGAGAATCCCCAGATGGCTTC	1892
Qy	1786	ATGGCCCCCAAAACATGGAGAGTCCCTGACAGATTATTACAGGTCCTCGGAGAGAAGCTAA	1845
Db	1893	ATGGCCCCCAAAACATGGAGAGTCCCTGACAGATTATTACAGGTCCTCGGAGAGAAGCTAA	1952
RESULT 2			
AAx84897			
ID	AAx84897 standard; cDNA to mRNA; 1875 BP.		
XX	AC	AAx84897;	
XX	24-SEP-1999	(first entry)	
DE	Human APS kinase/ATP sulphurylase gene.		
XX	KW	APS kinase/ATP sulphurylase; human; PAPS production;	
KW	3'-phosphoadenosine 5'-phosphosulphate; ds.		
XX	OS	Homo sapiens.	
XX	XX		
PN	JP11187883-A.		
XX	13-JUL-1999.		
XX	26-DEC-1997;	97JP-0360387.	
PF	XX		
XX	26-DEC-1997;	97JP-0360387.	
PR	XX		
XX	(NIRA) UNITIKA LTD.		
PA	(HUMA-) ZH HUMAN SCI SHINKO ZAIDAN.		
PA	WPI; 1999-451549/38.		
DR	P-PSDB; AAY22349.		
XX	New human-derived APS kinase/ATP sulfurylase gene - useful for		
PT	large scale preparation of 3'-phosphoadenosine 5'-phosphosulfate		
PT	(PAPS)		
PT			
XX	Claim 3; Page 7-8; 9pp; Japanese.		
CC	This sequence encodes the human-derived APS Kinase/ATP sulphurylase of		
CC	the invention. The enzyme may be used to prepare 3'-phosphoadenosine		
CC	5'-phosphosulphate (PAPS) on a large scale.		
XX			
SQ	Sequence 1875 BP; 526 A; 379 C; 499 G; 471 T; 0 other;		
Query Match 53.7%; Score 990; DB 20; Length 1875;			
Best Local Similarity 71.3%; Pred. No. 5.4e-295;			
Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;			
Qy	9	GATCAAGAGCAAAAGACGGAGAACCCAGCAGAGAATCCACCAATGTATGTATCAGGCCCA	68
Db	39	GAGCAATAACGCCGACAACTGGGGAATGCAGAGAGCAACCAATGTACCTACCAAGCCCA	98
Qy	69	CCATGTGACAGAGAAATAAGAGAGGGCAAGTGGTGGACAAAGGGTGGGTTCGAGGATG	128
Db	99	TCATGTGACAGAGAAACAGAGAGGTGAGTGGTGGGACCAAGAGGTGGCTTCGTGGTTG	158
Qy	129	TACCGTGTGGCTAACAGGTCTCTCTGGTGTCTGGAAACAAACGATAAGTTTTCGCCCTGA	188
Db	159	CACAGTTTGGCTTACAGGCTTGTCTGGAGCGGGAAGACTACTGTGAGCATGGCCCTTGA	218
Qy	189	GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGGAACAATGTCGGTCA	248

QY 609 CCACAGAGTGTGAACCTTTCGCAAGAGCAGAACTTGTACCCCTATATCTAATAAACA 668
Db 675 CCAGCAAGTGTGGAACCTTTCACAGGAAGCGGATATTGTACCTGTGGATGCTATGA 734
QY 669 TATCAGCAACTCTTTGTGCGGGAAGAAACAACCTTTCACCACTGCGAGCTGAGGCTGAAC 728
Db 735 AGTAAAGAACTATGTGCGGAGAAATAAATCTTATTTGGCAAAAACAGATGCGGAAC 794
QY 729 TCTCCCTTCATTATCAATTAAGCTGGAATCCAGTGGTCCAGGTTTTTGACGGAAG 788
Db 795 ATTACCACTGAAATTAATAAGTGGATATGAGTGGTGGAGTTTGGCAGAGG 854
QY 789 GTGGGCCACTCCCTCAAGGTTTTTCATGCGGGAAGAGTACTTACAGTTATGCACTT 848
Db 855 TTGGCAACCCCAATTAAGTGGTCTTATGAGAGAGGAGTACTTGCAGTCCCTTCATT 914
QY 849 TGACACCTGCTAGATGCGCTGATCAACATGAGATCCCACTTGTACTGCGCGCTC 908
Db 915 TGATTGCTCTGATGAGGTGCTAATTAAGTGGATATGAGTGGTGGAGTTTGGCAGAG 974
QY 909 TGCAGAGATGAAGACAGGCTGGAAGGTGAGCAAGTTTCTCTGCGACATGTTGGACG 968
Db 975 TCATGAGATGAAGAGAGGCTGGACGCTGTACAGCATTTGCTCTGATGATGAGGCG 1034
QY 969 GAGGTAGCTATCTTACGAGACGCTGAATCTTGAACACAGAAAGAGGAGCCTGTTC 1028
Db 1035 CCGTGTGGCCATCTTCGCAATCCAGATTTTTTTGAGCACAGGAAAGAGGCGCTGTGC 1094
QY 1029 CCGTGTGTGGGGACACATGTACAAACACCCCATATCAAAATGTTGATGGAAGTGG 1088
Db 1095 CAGACAGTGGGGAACGACATGCAAGACCCCATATTAAGATGTTGATGGAACAGG 1154
QY 1089 GGACTGGCTGTTGTTGGAGACCTTCAGGTGCTGGAGAAAATAAGATGGAATGATGGCT 1148
Db 1155 AGATTGGCTGATTGGAGGAGATCTCAAGTCTTGATGAGCTTTATTGGAATGATGCT 1214
QY 1149 GGACCAATACCGTCTGACACCTCTGGAGCTCAACACAGAAATGTAAAGAAATGAATGCTGA 1208
Db 1215 TGATCAGTATGCTTACTCTCTAGCTAAAGCAGAAATTTAAAGATATGAATGCTGA 1274
QY 1209 TGGGTGTTGCATTTCAGTTGCGCAATCTGTCACAAATGCGCAATGCGCTGTTGATGCA 1268
Db 1275 TGCTGTCTTGCAATTTCACTTACCAACCCAGTGCACAAATGAGATGCGCTGTTAATGCA 1334
QY 1269 GGACACCTGCGCAGGCTCTTAGAGAGGGGTACAAGCACCGGCTCTCTCTCTACTACACC 1328
Db 1335 GGATACCCATAAGCAACTTCTAGAGAGGGGTACCGGCGGCTCTCTCTCTCTCTCTCT 1394
QY 1329 TCTGGCGGCTGGACCAAGGATGAGATGCTGCTCTAGACTGGCGGATGAAGCAGCAGC 1388
Db 1395 TCTGGGTGGCTGGCAAGGATGACGATGTTCTCTTGTATGTTGGCTATGAAGCAGCATGC 1454
QY 1389 GGCTGTGCTGAGGAAGGGTCTCTGGATCCCAAGTCAACCAATGTTGCGCATCTTCCGTC 1448
Db 1455 TGCAGTGTGAGGAAGGAGTCTGAATCTTGAGACGACAGTGGTGGCCATCTTCCCATC 1514
QY 1449 TCCCATGTTATGCTGSCCCACAGAGGTCCAGTGGCACTGCAGGTCGCCGATGATTC 1508
Db 1515 TCCCATGATGATGCTGGAACCACTAGGTCAGTGGATGCGAGGACGAGGATGTTGC 1574
QY 1509 GGGTGCCAAATTTACATTTGGGGAGGAGCCCTGACAGGAATGCCCATCTCTGAAACCA 1568
Db 1575 AGAGAGCAACTTTTACATTTGTCAGCAGAGACCTCTGTCGATGCTCTCAGAAACAGG 1634
QY 1569 GAAGGATCTGATGAACCACTCATGCGGCAAGGCTTGGAGATGCGCCCTGCGCTCAC 1628
Db 1635 GAAGGATCTTATGAGCAAGTCAATGTTGGCAAGTGTGACGATGCGCCCTGTTTAAT 1694
QY 1629 CTCTGTGGAATCATCTCCAGTGGCTGCGCTTACAAAGCCAAAGCCAAAGCCATGGA 1688
Db 1695 CACTTTGGAATAGTTCCCTTTCAGTTGAGCTTACAAAGAAAGAGGCTATGGA 1754
QY 1689 CTTCTATGATCCAGAGGCAATGATGTTGACTTCTATCTCAGGAATCGAATGAGGAA 1748

Db 1755 CTACTATGACTCTGAACACCATGAAGACTTTTGAATTTATTTTCAGGAACACGAATGCGCAA 1814
QY 1749 GCTCGCGGGAAGAGAGAAATCCCCAGATGGCTTCATGCCCCCAAGCATGCAAGGT 1808
Db 1815 ACTTGTCTGGAAGAGGCCAGAAACCACTTGAAGGTTTTCATGCTCCCAAGGCTTGGACCGT 1874
QY 1809 CCTGACAGATTTATACAGGTCCTCGGAGAA 1838
Db 1875 GCTCAGAGATCTACAAATCCTTGGAGAA 1904
RESULT 4
AXX00717
ID AAX00717 standard; DNA; 2527 BP.
XX AC AAX00717;
XX 25-MAR-1999 (first entry)
XX Human secreted protein gene 76 clone HOSFD58.
DE Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX Homo sapiens.
OS
PN W09842738-Al.
XX 01-OCT-1998.
PD
XX 19-MAR-1998; 98WO-US05311.
PF
PR 30-MAY-1997; 97US-0050937.
PR 21-MAR-1997; 97US-0041276.
PR 21-MAR-1997; 97US-0041277.
PR 21-MAR-1997; 97US-0041281.
PR 21-MAR-1997; 97US-0042344.
PR 30-MAY-1997; 97US-0048069.
PR 30-MAY-1997; 97US-0048094.
PR 30-MAY-1997; 97US-0048095.
PR 30-MAY-1997; 97US-0048096.
PR 30-MAY-1997; 97US-0048099.
PR 30-MAY-1997; 97US-0048131.
PR 30-MAY-1997; 97US-0048135.
PR 30-MAY-1997; 97US-0048154.
PR 30-MAY-1997; 97US-0048160.
PR 30-MAY-1997; 97US-0048186.
PR 30-MAY-1997; 97US-0048187.
PR 30-MAY-1997; 97US-0048188.
PR 30-MAY-1997; 97US-0048350.
PR 30-MAY-1997; 97US-0048351.
PR 30-MAY-1997; 97US-0048352.
PR 30-MAY-1997; 97US-0048355.
PR 05-AUG-1997; 97US-0054804.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Brewer LA, Duan R, Ehner R, Ferrie AM, Florence KA;
PI Greene JM, Hu JS, Lafleur DW, Moore PA, NI J, Olsen HS;
PI Rosen CA, Ruben SM, Shi Y, Young P;
XX WPI: 1999-070066/06.
DR P-PSDB; ANW67913.
XX
PT New isolated human genes and the secreted polypeptides they encode -
useful for diagnosis and treatment of e.g. cancers, neurological

PT disorders, immune diseases, inflammation or blood disorders

PS Claim 1; Page 265-267; 385pp; English.

XX This sequence represents a nucleic acid molecule which encodes a secreted human protein. The gene number, and the clone it is derived from, are detailed in the descriptor line. The gene can be used to generate fusion proteins by linking to the gene to a human immunoglobulin Fc portion (e.g. AAX00602) for increasing the stability of the fused protein as compared to the human protein only.

CC The invention relates to 87 novel genes and their fragments (nucleic acid sequences: AAX00611-X00724; amino acid sequences AAW67807-W68004) which are useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. Also, pathological conditions can be diagnosed by determining the amount of the new polypeptides in a sample or by determining the presence of mutations in the new polynucleotides. Specific uses are described for each of the 87 polynucleotides, based on which tissues they are most highly expressed in (see AAX00611 for described uses).

XX SQ Sequence 2527 BP; 706 A; 508 C; 597 G; 716 T; 0 other;

Query Match 53.78; Score 990; DB 20; Length 2527;
Best Local Similarity 71.38; Pred. No. 6.4e-295;
Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;

QY	9	GATCAAGAACAAAGACGGAGAACACAGACAGAAATCCACCAATGTAGTCTATCAGGCCCA	68
DB	94	GAGCAATAACCGCGAGAACTGGGGAATGCAGAGAGCAACCAATGTCACTACCAAGCCCA	153
QY	69	CCATGTGAGCAGGATTAAGAGAGGGCAAGTGGTTGGAAAGGGGTTCGAGGATG	128
DB	154	TCATGTGAGCAGGAAAGAGAGGTGAGTGGTGGGACGAGAGGTGGCTTCGTGGTGG	213
QY	129	TACGCTGGCTTAACAGTCTCTCTGTGCTGTGAAACCAACGATAAGTTTGGCCCTGGA	188
DB	214	CACAGTTGGCTTAACAGGCTTGTCTGGAGCGGGAAAGACTACTGTGAGCATGGCCCTGGA	273
QY	189	GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAATGTCGTCA	248
DB	274	GGAGTACCTGTTGTCACTGATATTCATGCTACACTCTGATGGTGGTGCACAAATTCGTCA	333
QY	249	TGGCCTTAACAGAAATCTCGATCTCTCTGGGACAGAGAGAAATATCCGCGGAT	308
DB	334	AGGTCTCAATAAATACTTGGCTTTAGTCTGAAGACAGAGAGAAATGTTCCAGCGCAT	363
QY	309	TGCTGAGGTGGCTAAGCTGTTTGTCTGATGCTGTGCTGTGCTGCAATACCAAGCTTTATTC	368
DB	394	CGCAGAGTTGCTAAACTGTTTCAGATGCTGGCTTAGTGTGATCACAAGTTTCATATC	433
QY	369	TCCATTGCGAAAGATGCTGAGAAATGCCGCAAAATACATGAATCAGCGGGTGCAT	428
DB	454	ACCTTACACTCAGGATCGCAACAAATTAAGGTTTCACTGGGATCGATCTCAATATGA	513
QY	429	CTTTGAATATTGTAGATCAGCTCTTAATATTGTTGAAGCAGAGACGTAAGAGCGCT	488
DB	514	TTTGTGAAGTATTGTTGATGCTCCTCTGCAATGTTTGTGAACAGAGGATGTCAAAGGACT	573
QY	489	CTATAAAGGGCCAGAGCTGGGAGATTAAAGGATTACAGGTATTGATTCTGATTATGA	548
DB	574	CTACAAAAGCCCGGACGAGAAATTAAGGTTTCACTGGGATCGATTCTCAATATGA	633
QY	549	GAACCTGAAATCTCTGAGCGGTGCTTAAACCAATTTGTCACAGTGAAGTGTGT	608
DB	634	AAAGCCAGAGGCGCCCTGAGTTGGTGTGTAAGAACAGACTCCTGTGATGTAATGACTGTGT	693
QY	609	CCACCAAGTGTGGAATCTCTGCAAGACGAGACATTGTACCCCTATATAATCAAGA	668
DB	694	CCAGCAAGTTGTGGAATCTCTACAGGAACGGGATATTGTACCTGTGGATGCACTCTATGA	753
QY	669	TATCCAGCAACTCTTTGTGCGGAAACAACTTGTGACCGCTCCGAGCTGAGGCTGAAAC	728
DB	754	AGTAAAGAACTATATGTGCCAGAAATAAATTCATTTGTCGCAAAACAGATCGGCAAC	813

QY	729	TCCTCCCTTCATTATTAAGCTGGATCTCCAGTGGGTCCAGGTTTGTAGCGAAGG	788
DB	814	ATTACCAGCACTGAAATTAATAAGTGGATATGCAGTGGGTGGAGGTTTGGCAGAAGG	873
QY	789	CTGGGCACTCCCTCAAGGTTTCAATGCGGAGAGAGTACTTACAGGTTATGCACTT	848
DB	874	TTGGCAACCCCATTTGAATGGCTTTATGAGAGAGAGGAGTACTTGCAGTGGCTTCAAT	933
QY	849	TGACACCCCTCTAGATCATCGCTGATCAACATGAGCATCCCCATTTGCTACTGCCCCCTC	908
DB	934	TGATTGCTTCTGGATGGAGGTGTCAATTAACCTTGTTCAGTACCTATAGTTCTGCACTGCGAC	993
QY	909	TGACAGAGATAAGACACGCTGGAAGGTTGCAGCACTTGTCTGTCACATGTTGGGACG	968
DB	994	TCATGAAGATAAGAGAGGCTGCAGCGCTGTACAGCATTTGCTGTGATGTATGAGGCGG	1053
QY	969	GAGGTAGCTATCTTACGAGACGCTGAATTTCTATGAACAGAAAGAGAGAGCTGCTTC	1028
DB	1054	CCGTGTGGCCATTTCTGCAATCCAGAGTTTGTGACACAGGAAAGAGGCGCTGTGC	1113
QY	1029	CCGTGTTGGGGACACATGTCAAAACACCCCATATCAAAATGTTGATGAAAGTGG	1088
DB	1114	CAGACAGTGGGAAACGACATGCAAGAACCCCTATATTAAAGATGTTGATGGAACAAGG	1173
QY	1089	GGACTGGCTGGTTGGTGAGACCTTCAGGTGCTGGAGAAATTAAGATGGAATGATGGCT	1148
DB	1174	AGATTGGCTGATTGGAGGAGATCTCAAGTCTTGGATCGAGTTATTGGAATGATGCTCT	1233
QY	1149	GGACCAATACCTGCTGACACCTCTGGAGCTCAACAGAAATGTAAAGAAATGATGCTGA	1208
DB	1234	TGATCAGTATGCTTACTCTCTACTGAGCTAAGACAGAAATTTAAAGATGATGAATGCTGA	1293
QY	1209	TGCGGTGTTTGGCAATCCAGTTGCGCAATCTGCTCCCAATGGCCTGTTGATGCA	1268
DB	1294	TGCTGCTTGGCATTTCACTACGCAACCCAGTGCACATGGACATGCCCTGTTAATGCA	1353
QY	1269	GGACACTGCCGAGGCTCTTAGAGAGGGCTTACAAGCACCCGGTCTCTTACTATACACC	1328
DB	1354	GGATACCCATAAGCAACTTCTAGAGAGGGCTACCGCGCCCTGCTCTCTCTCCACCC	1413
QY	1329	TCTGGGCGGCTGACCAAGATCAGCTGCTCTAGACTGGCGGATGAAGCAGCACGC	1388
DB	1414	TCTGGGTGGCTGACAAAGGATGACGATGCTCTTGTATGTGGCGTATGAAGCAGCATGC	1473
QY	1389	GGCTGTCTCGAGAGAGGGTCTCTGGATCCCAAGTCAACCATTTGTTGCCATCTTTCGCTC	1448
DB	1474	TGCAGTGTGCGAGAGGAGTCTGATCCTGAGACGACAGTGTGGCCATCTTCCCATC	1533
QY	1449	TCCCATGTTATATGCTGGGCCCCACAGAGTCCAGTGGCACTGCGAGTCCCGGATGATGC	1508
DB	1534	TCCCATGATGATGCTGGACCAACTGAGGTCCAGTGGCATTTGCAGAGCAGGATGGTTGC	1593
QY	1509	GGGTGCCAATTTCTACATTTGGGAGGAGCCCTGAGGAATCCCATCTCTCAAAACCA	1568
DB	1594	AGAGCCCACTTTTACATTTGTCAGAGACCCCTGCTGGCATGCTCTTCCATCCAGAAACAGG	1653
QY	1569	GAAGGATCTGATGAACCCACTCATGGGCAAGGCTTTCAGCATGCGCCCTGGCCCTCAC	1628
DB	1654	GAAGGATCTTTATGACCAAGTCTGTCGCAAGTGTGACCATGGCCCTGTTTAAAT	1713
QY	1629	CTCTGTGGAATCATTTCCATTCGAGTGGCTGCTTACACAAAGCCCAAAAGCCATGGA	1688
DB	1714	CACCTTTGGAATAGTTTCCCTTTCAGTTGAGTTCAGCTTACAACAGAAAAGAGCGTATGA	1773
QY	1689	CTTCTATGATCCAGCAAGGCACAAATGATTTGACTTTCATCTCAGGAACCTGAAATGAGAA	1748
DB	1774	CTACTATGACTTGAACACCATGTAAGACTTTGAATTTATTTTCAAGAACAGATGCGCAA	1833
QY	1749	GCTCGCGGAGAGAGAAATCCCGAGATGGCTTTCATGGCCCCCAAGACATGGAAGGT	1808
DB	1834	ACTTGCTCGAGAAGGCCAGAAACCACTGAAGGTTTTCATGGCTTCCCAAGGCTTGGACCGT	1893

Db 874 TTGGCAACCCCAATGAATGGCTTTATGAGAGAGAGGGAGTACTTGCAGTGCCTTCATTT 933
 QY 849 TGACACCTCTGCTAGATGATGGCTGATCAACATGAGCATCCCATTTGACTGCCCTCTC 908
 Db 934 TGATTGCTCTCTGGATGGAGGTGTCATTAACCTTGTCTAGTACCTATATGTTCTGACTCGGAC 993
 QY 909 TGCAGAGGATAAGACACGGCTGAAGGGTGCAGCAAGTTTGTCTCTGGCACATGGTGGAGC 968
 Db 994 TCATGAAGATAAAGAGAGGCTGGACGGCTGTACAGCATTTGCTCTGATGTATGAGGGCGC 1053
 QY 969 GAGGTAGCTATCTTACGAGACGCTCAATCTATGAACACAGAAAGAGGACGCTGTTTC 1028
 Db 1054 CCGTGGCCATTTCTCGCAATCCAGAGTTTTTGAGCACAGGAAGAGGAGCGCTGTC 1113
 QY 1029 CCGTGTGGGGGACAAACATGTACAAAACACCCCATATCAAAATGGTGTGAAAGTGG 1088
 Db 1114 CAGACAGTGGGAACGACATGCAAGAACCCCTATATTAAGATGGTGTGAAACAGG 1173
 QY 1089 GGAAGTGGTGGTGGAGACCTTCAGGTGTGGGAGAAAATAAGATGAATGAGGTGGGT 1148
 Db 1174 AGATTGGCTGATTTGGAGAGATCTTCAAGTCTTGGATCGAGTTTATTTGAATGATGGTCT 1233
 QY 1149 GGACCAATACCGCTCTCACACCTCTGAGCTCAACACAGAAATGTAAGAAATGAATGCTGA 1208
 Db 1234 TGATCAGTATCGTCTTACTCTACTGAGCTTAAGCAGAAATTTAAAGATATGAATGCTGA 1293
 QY 1209 TGGGTGTTTGGCATTCAGTTCGCAATCTGTCTCCCAATGGCCATGCCCTGTGGATGCA 1268
 Db 1294 TGCTGTCTTGGCATTTCAACTAGCAACCCAGTGCACAAATGGACATGCCCTGTGTAATGCA 1353
 QY 1269 GNAACCTCGGAGGCTCTAGAGAGGGCTTACAGCACCCGGTCTCTCTACTACTACACC 1328
 Db 1354 GGATACCCCAATGAACAACTTCTAGAGAGGGCTACCGCGCCCTGTCTCTCTCTCCACCC 1413
 QY 1329 TCTGGCGCTGGACCAAGATGACATGCTGCTCTAGACTGCGGATGAAGCAGCACGC 1388
 Db 1414 TCTGGTGGCTGGCAAGAGATGACATGTTCTCTGTGCTGGCTATGAGCAGAGTGC 1473
 QY 1389 GCGTGTGCTGGAGAGGGTCTGTGATCCCAAGTCAACCATTTGTGCCATCTTTCCTGTC 1448
 Db 1474 TGCAGTGTGGAGAGAGGATTTCTGAATCTGAGAGCAGAGTGTGGCCATCTTCCCATC 1533
 QY 1449 TCCTATGTTATGCTGGCCACACAGAGCTCCAGTGGCACTGAGTCCCGGATGATTCG 1508
 Db 1534 TCCATGATGATGCTGGACCACTGAGGTCCAGTGGCATTTGCAGAGCAGCGATGTTGC 1593
 QY 1509 GGTGTCCAATTTCTACATTTGGGGAGGACCTCGAGGAATGCCCATCTCTGAAACCAA 1568
 Db 1594 AGGAGCAACTTTTACATTTGAGAGAGACCTGCTGGCATGCTCATCCAGAACAGG 1653
 QY 1569 GAAGATCTGTATGAACCACTCATGGGGCAAGGTCTTTGAGCATGGCCCCCTGGCCTCAC 1628
 Db 1654 GAAGGATCTTTATGAGCAAGTCTGTCGCAAGTGTCTGAGGATGGCCCTGGTGTAAAT 1713
 QY 1629 CTCTGTGGAATCATTCATTCCTGAGTGGCTGCTACAAACAGCCAAAGAGCCATGCA 1688
 Db 1714 CACTTGGAAATAGTTTCCCTTTGAGTTGAGCTTACAAACAGAAAGAGCGTATGGA 1773
 QY 1689 CTCTATGATCCAGCAAGGCACATGAGTTTGAATCTCTCAGGAACCTCGAATGAGGAA 1748
 Db 1774 CTACTATGACTCTGAACACCATGAGACTTTGAATTTATTCAGGAACAGATGCGCA 1833
 QY 1749 GCTGCCCCGGAGGAGAGAAATCCCCCATGCTGCTGAGTGGCCCCCAAGAGATGGAAGT 1808
 Db 1834 ACTTGTCTGAGAGGGCCAGAAACCACTGAAGGTTTCATGGTCTCCCAAGGCTTGGACCGT 1893
 QY 1809 CTTGACAGATTTATCAGGCTCCCTGGAGAA 1838
 Db 1894 GCTGACAGAATACTACAAATCTTTGGAGAA 1923

ID AAS94855 standard; DNA; 2646 BP.
 XX AAS94855;
 AC
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Human DNA sequence #110 expressed during foam cell differentiation.
 XX Human; foam cell differentiation; atherosclerosis; cerebral stroke;
 KW cardiovascular disorder; coronary artery disease; gene therapy; ds.
 XX Homo sapiens.
 OS
 XX WO200177389-A2.
 PN
 XX 18-OCT-2001.
 PD
 XX
 PF 04-APR-2001; 2001WO-US11128.
 XX
 PR 05-APR-2000; 2000US-195106P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Shiffman D, Somogyi R, Lawn R, Sellhamer JJ, Porter GJ, Mikita T;
 PI Tai J;
 XX
 DR WPI; 2002-010925/01.
 XX
 CC Composition useful for diagnosis of conditions, disorders or diseases
 CC associated with atherosclerosis, comprises several polynucleotides that
 CC are differentially expressed in foam cell development -
 XX
 CC Claim 1; Page 162-163; 315pp; English.
 XX
 CC The present invention relates to the isolation of human polynucleotide
 CC sequences that are differentially expressed during foam cell
 CC differentiation. The polynucleotide sequences of the invention or a
 CC composition comprising these polynucleotides are useful as a high
 CC throughput method for detecting altered expression of one or more
 CC polynucleotides in a sample. The polynucleotides can be used in the
 CC diagnosis of disorders associated with foam cell development such as
 CC atherosclerosis, cerebral stroke, and cardiovascular disorders such as
 CC coronary artery disease. The polynucleotide sequences can also be used
 CC as PCR primers and probes. The polynucleotides of the invention are also
 CC useful in gene therapy. AAS94746-AAS95021 represent the human
 CC polynucleotide sequences of the invention which are differentially
 CC expressed during foam cell differentiation.
 XX
 SQ Sequence 2646 BP; 715 A; 566 C; 634 G; 727 T; 4 other;
 Query Match 53.7%; Score 990; DB 24; Length 2646;
 Best Local Similarity 71.3%; Pred. No. 6.6e-295;
 Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;
 QY 9 GATCAGAGCAAAAGACGAGGAGGCAAGTGGTGAACAAGGGTGGTTCGAGGATG 68
 Db 220 GAGCAATAACGCGCAACTGGGGAATGAGAGAGCAACCAATGTCCATACCAAGCCCA 279
 QY 69 CCATGTGAGCAGGAATAAGAGAGGCAAGTGGTGGTGAACAAGGGTGGTTCGAGGATG 128
 Db 280 TCATGTGACAGCAACAGAGAGGTCAGGTGGTGGGACCAAGAGGTGGCTTTCGTGTTG 339
 QY 129 TACCGTGTGGCTTAACAGGTCTCTCTGGTCTGCTGAAACAAACAGATAAGTTTGGCCCTGGA 188
 Db 340 CACAGTTTGGCTTAACAGGTCTTGTCTGGAGCGGAAAGACTACTGTGAGCATGGCTTGA 399
 QY 189 GGAGTACCTTGTCTCCCATGCCATCCCTTTGTACTCCCTGGATGGGACAAATGTCGTC 248
 Db 400 GGAGTACCTTGGTGTTCATGTTCCATGCTGCTACACTCTGAGTGGTGAACAATTCGTCA 459
 QY 249 TGGCCTTTAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGAAATATCCGCCGGAT 308
 Db 460 AGGTCTCAATAAATCTTGGCTTTAGTCTCTGAACACAGAGAGAGAAATGTTCCGAGCAT 519

QY	1389	GGCTGTGCTCGAGGAGGGTCTCGATCCCAAGTCAACCATTTGTTGCCATCTTTCCGTC	1448
Db	1600	TGCAGTGTTTGGAGGAGGAGTCTCGAATCTCGACGACAGTGGTGCCCATCTTCCCATC	1659
QY	1449	TCCCATGTTATATGCTGGCCCCACAGAGGTCCAGTGGCACTGCAGGTCCCGGATGATTGC	1508
Db	1660	TCCCATGATGATGCTGGACCACTGAGGTCCAGTGGCATTTGCAGACACGGATGGTTGC	1719
QY	1509	GGGTGCCAATTTCTACATTTGTGGGGAGGGACCTCGAGGAATGCCCATCCTGAAACCA	1568
Db	1720	AGAGGCCAACTTTTACATTTGTGGAGGAGACCTGTGTGGCATGCTTCATCCAGAAACAGG	1779
QY	1569	GAAGGATCTGATGAACCCACTCATGGGGCAGGTCTTCAGCATGCCCCCTGSCCTCAC	1628
Db	1780	GAAGGATCTTTATGAGCCAAGTCATGGTGCCAAGTGTGACCATGGCCCTGTTTAAAT	1839
QY	1629	CTCTGTGGAATCATTTCCATTCGAGGTGGCTGCCTACAAACAAAGCCAAAAAGCCATGGA	1688
Db	1840	CACTTTGGAAATAGTTCCCTTTCCAGTTGCGAGTTTACAACAAGAAAAAGCGTATGGA	1899
QY	1689	CTTCTATGATCCAGCAAGGCACAAATGATTTTGTACTTCATCTCAGGAACCTCGATGAGAA	1748
Db	1900	CTACTATGACTCTGAAACACCATGAAGACTTTGAAATTTATTTTCAGGAACAGATGGCCAA	1959
QY	1749	GCTCGCCCCGGAAGGAGAGAAATCCCCAGATGGCTTCATGGCCCCCAAAAGCATGGAAGGT	1808
Db	1960	ACTTGTCTGAGAGGCCAGAAACCACTGAAGGTTTCATGGCTCCCAAGGCTTGGACCGT	2019
QY	1809	CCTGACAGATTTATACAGGTCCTCGGAGAA	1838
Db	2020	GCTGCACAGAATACTACAAATCCTTTGGAGAA	2049

RESULT 7

[illegible]

Qy	1610	GCATGGCCCTGGCCTCACC-----	1629
Db	1615	AGATGCCCGAGGACTAGACAGCATGGAGGTAAGGTAACCACTGCTTAATCGGCCAAATC	1556
Qy	1630	-----TCTGTGGAATCATTCCTCATTCGAGTGGCTGCCTAC	1665
Db	1555	TAGCCGNATGAATCGATATTTCTTTTAGATCCTGCCCTTCGTGTGCTGCCTAT	1496
Qy	1666	AACAAAGCCAAAAGCCATGGACTTCTATGATCCAGCAAGGCACAAATGAGTTTGACTTC	1725
Db	1495	GACAAGAGCGCCAGTAGAATGGCTTCTTCGAGCCCAAGAGGAAGGATGAGTTTGAGTTC	1436
Qy	1726	ATCTCAGGAAGCTCGATGAGGAAGCTCGCCCGGAAGGAGAGAGAAATCCCCAGATGGCTTC	1785
Db	1435	ATCTCGGGGAACCAAGATGGCACCTTGGCCAAACGGGAGCCAGTCCGCCGATGGCTTC	1376
Qy	1786	ATGCGCCCCAAACGATGGGAAGGTCTCTGACAGATTATACAGGTCCCTGGAGAGAAGAACTAA	1845
Db	1375	ATGGAACCGGAGGCTTGGCGCATCTCTGGCCACCTACTACAGAACTGCCGCAGTCGTAA	1316
RESULT 9			
AAT04849			
ID	AAT04849 standard; cDNA to mRNA; 990 BP.		
AC	AAT04849;		
XX			
DT	28-MAY-1996 (first entry)		
XX			
DE	Human adenosine 5' phosphosulphate kinase 990 bp gene fragment.		
XX			
KW	Phosphosulphate kinase; genetic disease; diagnosis; probe; ds.		
XX			
OS	Homo sapiens.		
XX			
PN	JP07250687-A.		
XX			
PD	03-OCT-1995.		
XX			
PF	14-MAR-1994; 94JP-0069861.		
XX			
PR	14-MAR-1994; 94JP-0069861.		
XX			
PA	(NIRA) UNITIKA LTD.		
XX			
DR	WPI; 1995-370475/48.		
XX			
PT	DNA fragment of human adenosine 5' phospho:sulphate kinase - used as		
PT	a probe in the detection of genetic disease and for the production		
PT	of 5' phospho:sulphate kinase		
XX			
PS	Claim 1; Page 6; 6pp; Japanese.		
XX			
CC	AAT04849 and AAT04850 are fragments of the human adenosine 5'		
CC	phosphosulphate kinase (5' PSK) of 990 and 600 bp respectively.		
CC	may be used as probes for the detection of genetic disease. The 990		
CC	bp fragment may also be used for low cost prodn. of human 5' PSK.		
XX			
SQ	Sequence 990 BP; 281 A; 182 C; 267 G; 260 T; 0 other:		

[illegible]

Qy	156	TCCTGGAAACAACACGATAAGTTTGGCCCTGGAGGAGTACCTTGTCTCCCATGCCATCC	215
Db	123	AGCGGGAAGACTACTGTGAGCATGGCCCTGGAGGAGTACCTGTTGTTCATGSGTATCC	182
Qy	216	TTGTTACTCCCTGGATGGGACAATGTCGGTCATCGCTTAACAGAAATCTCGGATTC	275
Db	183	ATGCTACACTCTGGATGSGACAATATTCGTCAGGCTCTCAATAAANAATCTTGGCTTTAG	242
Qy	276	TCCTGGGACACAGAGGAGAAATATCCGCCGGAATGCTGAGGTGGCTAAAGCTGTTTGTCTGA	335
Db	243	TCCTGGAAGACAGAGAAGAGAATGTTCGACGGATCGCAGAAGTGTCTAAACTGTTTGCAGA	302
Qy	336	TGCTGGTCTGCTCTGCAATACCAAGCTTTATTTCTCCATTTCGAAGGATCGTGAAGATCC	395
Db	303	TGCTGGCTTAGTGTGCATCAAGAATTCATACCTTACACTCAGGATCGCAACAATGC	362
Qy	396	CCGCAAAATACATGAATCAGCAGGGCTGCCATTCCTTTGAAATATTTGTAGATGCACCTCT	455
Db	363	AAGCAAAATTCATGAAGGTGCAAGTTTACCGTTTTTTCGAAGTATTTTGTGATGCTCCTCT	422
Qy	456	AAATATTTGTGAAGACAGACGATGAAGAGGCTCTATAAAGGGCCAGCTGGGGAGAT	515
Db	423	GCATGTTTGTGAACAGAGGGATGTCAAAGSACTCTACAAAAGCCCGGCGAGSAGAAAT	482
Qy	516	TAAAGGATTACAGGTAATTGATTCCTGATTATGAGAAACCTGAAACCTCTGAGCGTGTGCT	575
Db	483	TAAAGTTTCACTGGGATCGATTCTGAATATGAAGACGAGGCCCTGAGTTGGTGCT	542
Qy	576	TAAACCAATTTGTCCACAGTGAAGTGTGTCCACAGGTAGTGAACCTTCTGCACAGA	635
Db	543	GAAACAGACTCCGTGTATGTAATGACTGTGTCCAGCAAGTTGTGGAACCTCTACAGGA	602
Qy	636	GCAGAACATGTACCCCTATACTATANTCAAGAGATCCACAGCACTTTTGTGCCGGAANA	695
Db	603	ACGGGATATTGACTCTGGATGCATCTTATGAAGTAAAGAGCACTATATGTGCCAGAAA	662
Qy	696	CAAACTTGACCACGCTCGGAGCTGAGGCTGAAACTCTCCCTTCATTATCAATTAAGCT	755
Db	663	TAAACTTCAITTTGGCAAAAACAGATCGGAAACATTAACACAGTGAATAATATAAGT	722
Qy	756	GGATCTCCAGTGGGTCCAGGTTTGGCGAAGGCTGGGCCACTCCCTCAAAGGTTTCAT	815
Db	723	GGATTCGAGTGGGTGCAGGTTTGGCAGAAGGTTGGGCAACCCCATTTGAATGGCTTTAT	782
Qy	816	GCGGGAGAAGGAGTACTACAGGTTATGCACTTTGCACACCTGCTAGATGATGGCGTGAT	875
Db	783	GAGAGACAGGAGTACTTGCAGTCCCTTCATTTTGTGTTCTCTGGATGGAGGTGTCAT	842
Qy	876	CAACATCAGCATCCCCATTTGACTGCCGCTCTCTGCAGAGAGATTAAGACACGGCTGGAAG	935
Db	843	TAACTTTCAGTACCTATAGTTCGTGACTGGCACTCATGAAGATAAAGAGAGGCTGGATGG	902
Qy	936	GTGCAGCAAGTTTGTCTCTGGCACAATGGTGGACGGAGGTAAGCTATCTTACGAGACGCTGA	995
Db	903	CTGTACAGCATTTGCTCTGATGTAT - GAGGCCGCCGTGTGCCCATCTCTCGCAATCCAGA	961
Qy	996	ATTCTATGAACACAGAAAGAGGAACGC	1023
Db	962	GTTTTTTGGACACAGGAAGAGAGGAGCC	989

RESULT 10	
AAC34705	
ID	AAC34705 standard; DNA; 1695 BP.
XX	
XX	
AAC34705;	
XX	
XX	
DT	17-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 7588.
XX	
XX	
KW	Hybridisation assay; genetic mapping; gene expression control;
KW	protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 06-MAY-1999; 99US-0132487.
XX 07-MAY-1999; 99US-0132863.
XX 11-MAY-1999; 99US-0132456.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 14-MAY-1999; 99US-0134370.
XX 18-MAY-1999; 99US-0134768.
XX 19-MAY-1999; 99US-0134941.
XX 20-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139750.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR	10-SEP-1999;	99US-0153070.	
PR	13-SEP-1999;	99US-0153758.	
PR	15-SEP-1999;	99US-0154018.	
PR	16-SEP-1999;	99US-0154039.	
PR	20-SEP-1999;	99US-0154779.	
PR	22-SEP-1999;	99US-0155139.	
PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	21-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 20.8%; Score 384; DB 21; Length 1695; Best Local Similarity 59.6%; Pred. No. 1.3e-107; Matches 705; Conservative 0; Mismatches 465; Indels 12; Gaps 3;			
QY	676	GAACCTCTTGGCCGGAACAACTTGACCACGTCGAGCTGAGGCTGAACTCTCCCT. 735	
Db	312	GAATTGATTGACCGAACTGAGATTGGAGTGAAGAAAGCTGAATCAGAGACGATCCCT 371	
QY	736	TCATTATCAATTACTAAGTGGATCTCCAGTGGCTCCAGGTTTGGAGCGAAGGCTGGCC 795	
Db	372	AAAGTGAAGCTGAATCAGATTGATCTTGAGTGGGTTCTATGATGATTAGTGAAGTGGCT 431	
QY	796	ACTCCCTCAAAGTTTTCATCGGGAGAGAGTACTTACAGGTTATGCACTTTTGACACC 855	
Db	432	AGTCCTTTAAAAGGTTTCATGAGAGAAGATGAGTATTTGCAAAAGTTTGCATTTAATCT 491	
QY	856	C-----TGCTAGATGATGGGTGATCAACATAGAGATCCCATTTGCTCCCTCTCT 909	
Db	492	CTTAGGTTAAAAAATGGAGCTTTTGTGAACATGCTCTCTCTCTATTTGTTGTTGAT 551	
QY	910	GCAGAGATGAAGACACGCTGGAAGGCTGCAGCAAGTTTGTCTCTGG--CACATGGTGA 966	
Db	552	GATGATCTAAGGACAGATCGGATCATCTGAGAAATGTTGCTCTGTTTGCCTCAAGGA 611	
QY	967	CGGAGGTAGCTATCTTACGAGACGCTGAATTTCTATGAACACAGAAAGAGACGCTGT 1026	
Db	612	GATATCATTTGTTCTCTCCGAAGTGTGGAGATTTTACAACACACATAAGGAAGGATA 671	

QY	1027	TCCCGTGTGGGGGACACATGTACAAAACACCCCATAT---CAAAATGGTGATGAA 1083	
Db	672	GCTAGAACTGGGGAACACTTCTCCAGGATTGCTTATGGAAGAAATATATTACTCA 731	
QY	1084	AGTGGGACTGGCTGGTGGAGACCTTCAGGTGGAGAGAAATAGATGGAATGAT 1143	
Db	732	TCTGGAACACTGGCTGATTGGTGGTATTAGAAATTTTGAACCGATTAAAGTATAACGAT 791	
QY	1144	GGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACACAGAAATGTAAAGAAATGAAT 1203	
Db	792	GGCTTGATCATATTACAGGCTCTCTCTAAACAGCTTAGGAGAGATTTGATTAATCGTCAA 851	
QY	1204	GCTGATGGGTGTTTGGCATTCAGTTCGCAATCCTGTCCAAATGGCCATGCCCTGTG 1263	
Db	852	GCAGATGCTGTTTTTGGCTTTTTCAGTTAAGGAATCTGTGCAATAGGACATGCTCTGTG 911	
QY	1264	ATGAGGACACCTGCCGAGGCTCCTAGAGAGGGCTACAAGCACCCCGTCTCTCTACTA 1323	
Db	912	ATGAATGATACACGAAAGGCTTTTGGAAATGGGGTACAAGAACCCCTGTTCTCTGCTT 971	
QY	1324	CACCTCTGGCGGCTGGACCAAGGATGACGATGTCCTCTAGACTGGCGGATGAAGCAG 1383	
Db	972	CATCTTTAGGAGGTTTCACTAAAGCTGATGATGTTCTCTCTGATTTAGATGAACAA 1031	
QY	1384	CACGCGCTGTGCTCGAGGAAGGGTCTCGATCCCAAGTCAACCAATTTGTCCTATCTTT 1443	
Db	1032	CATAGCAAGGTTCTAGAAGATGGAGTTCTTGACCGGAAGACTACTATTGTCGATCTTT 1091	
QY	1444	CCGTCTCCCATGTTATATGCTGGCCCAACAGAGTCCAGTGGCAGTGGCGGATG 1503	
Db	1092	CCGTACCGGATGCACTATGCTGGTCCCAACTCAAGTTCAATGGCAGCCCAAGCAAGGAT 1151	
QY	1504	ATTGCGGTGCAATTTTCTACATTTGTTGGGAGGAGCCCTGCAGGAATGCECCATCTGAA 1563	
Db	1152	AACGCTGCTGCTAATTTCTACATTTAGTCTGATCTCTGAGGAATGGCCACCCCTACT 1211	
QY	1564	ACCAAGAGGATCTGTATGAACCCACTCATCGGGGCAAGGCTTTGAGCATGCCCTGGC 1623	
Db	1212	GAGAAGAGAGATCTATATGATCTCTGTCGCAAAAGAGTTTCTTTCAITTTCTGGAACAAAGATG 1271	
QY	1624	CTCACCTCTGTGGAATCATTCCTATTCGAGTGGCTGCTTACACAAAGCCCAAAAGGCC 1683	
Db	1272	CTTGAGAACTCAATATTTCTCCATTTAGGTTTACAGTTTACGATACATCGAGNAGAG 1331	
QY	1684	ATGACTTCTATGATCCAGCAAGCAATGAGTTTGAATTTCACTTCACTCAGGAACCTGAAATG 1743	
Db	1332	ATGCGTTTTTTCGATCTCTCTGTCGCAAAAGAGTTTCTTTCAITTTCTGGAACAAAGATG 1391	
QY	1744	AGGAAGCTGCCCGGAGGAGAGAAATCCCCACAGATGGCTTCATGGCCCCCAAGACATGG 1803	
Db	1392	CGAACTTATGCAAGAACAGGGGAGAAATCTCTCCAGATGGATTTATGTGTCGAGTGGATGG 1451	
QY	1804	AAGTCTCTGACAGATTATTACAGCTCCCTGGAGAGAACTAA 1845	
Db	1452	AATGTTCTTCAATACTATGAGAGCTTGCAGAGAGAGAA 1493	

RESULT 11
 AAA51055
 ID AAA51055 standard; DNA; 3684 BP.
 XX
 AC AAA51055;
 XX
 DT 09-OCT-2000 (first entry)
 XX
 DE SAMS promoter::ATPS DNA fragment.
 XX
 KW S-adenosyl-L-methionine synthetase; SAMS; probe; promoter; embryo;
 KW constitutive; tissue-specific; development-specific; ATP synthase;
 XX ATPs; herbicide resistance; pathogen resistance; ss.
 OS Glycine max.
 OS Synthetic.


```
Db 61 GACACTCCCGCAGCTCCTAGAGAGGGCTACAAAGCCCGGTCTCTACTACACGCT 120
QY 1330 CTGGCGGCTGGACCAAGATGACGATGCTCTAGACTGGCGGATGAAGCAGCACCGG 1389
Db 121 CTGGCGGCTGGACCAAGATGACGATGCTCTAGACTGGCGGATGAAGCAGCACCGG 180
QY 1390 GCTGTGCTCGAGGAAGGGTCTGTGATCCCAAGTCAACCATTTGTCATCTTCCGCT 1449
Db 181 GCTGTGCTCGAGGAGGGTCTGTGATCCCAAGTCAACCATTTGTCATCTTCCGCT 240
QY 1450 CCCATGTTATATGCTGGCCACACAGAGTCCAGTGGCACTGCAGTCCCGGATGATTCGG 1509
Db 241 CCCATGTTATATGCTGGCCACACAGAGTCCAGTGGCACTGCAGTCCCGGATGATTCGG 300
QY 1510 GGTGCCAATTCTACATTGTGGGAGGACCTGCAGGAATGCCCATCTGAAACCAAG 1569
Db 301 GGTGCCAATTCTACATTGTGGGA-GGACCTGCAGGAATGCCCATCTGAAACCAAG 359
QY 1570 AAGGATCTGTATGAACCCACTCATGTTGGGCAAGGCTTTGAGC 1611
Db 360 AAGGATCTGTATGAACCCACTCATGTTGGGCAAGGCTTTGAGC 401

RESULT 14
AAC47786
ID AAC47786 standard; DNA; 1687 BP.
XX AC AAC47786;
XX XX
DT 18-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 55114.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX PD
XX PD
XX PD
PF 25-FEB-2000; 2000EP-0301439.
XX XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125789.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
```



```
Db 1149 GAGAAACGTGATCTTTACGATCTGATCATGGAAGAAAGTACTAAGCATGCGCAGGA 1208
QY 1624 CTCACCTCTGTGGAATCATTCATTCCGAGTGGCTCTACACAAAGCCAAAAGCC 1683
Db 1209 CTCGACGACACTCAACATCTCTTCAGGTGCTGATATGACAAAGCGCAAGCGAAG 1268
QY 1684 ATFGACTTCTATGATCCAGCAAGGACAAATGAGTTTGACTTCATCTCAGGAACCTGAAATG 1743
Db 1269 ATGGCTTCTTCGATCCCTCGAGGCTCAAGATTCTTGTTCATCTCCGGCACTAAGATG 1328
QY 1744 AGGAAGCTCGCCCGGAGGAGAGAGATCCCGCAGATGGCTTCATGGCCCCCAAGCGATGG 1803
Db 1329 CGCATTGGCAAGAAACAAACGAAACCCGCGAGACGGTTTATGTGCCCGAGGTGGATGG 1388
QY 1804 AAGGTCTCTGACAGATTTATACAGTCCCTG 1833
Db 1389 AAGGTCTCTGGTACTATGAGAGCTTG 1418

RESULT 15
AAC42592
ID AAC42592 standard; DNA; 1533 BP.
XX
AC AAC42592;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 36133.
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 28-APR-1999; 99US-0130891.
XX 30-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 04-MAY-1999; 99US-0132407.
XX 05-MAY-1999; 99US-0132484.
XX 06-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.
XX 07-MAY-1999; 99US-0132487.
XX 11-MAY-1999; 99US-0132863.
XX 14-MAY-1999; 99US-0134256.
XX 14-MAY-1999; 99US-0134218.
XX 14-MAY-1999; 99US-0134219.
XX 14-MAY-1999; 99US-0134221.
XX 18-MAY-1999; 99US-0134370.
XX 19-MAY-1999; 99US-0134768.
XX 20-MAY-1999; 99US-0134941.
XX 21-MAY-1999; 99US-0135124.
XX 21-MAY-1999; 99US-0135353.
XX 24-MAY-1999; 99US-0135629.
XX 25-MAY-1999; 99US-0136021.
XX 27-MAY-1999; 99US-0136392.
XX 28-MAY-1999; 99US-0136782.
XX 01-JUN-1999; 99US-0137222.
XX 03-JUN-1999; 99US-0137528.
XX 04-JUN-1999; 99US-0137502.
XX 07-JUN-1999; 99US-0137724.
XX 08-JUN-1999; 99US-0138094.
XX 10-JUN-1999; 99US-0138540.
XX 10-JUN-1999; 99US-0138847.
XX 14-JUN-1999; 99US-0139119.
XX 16-JUN-1999; 99US-0139452.
XX 16-JUN-1999; 99US-0139453.
XX 17-JUN-1999; 99US-0139492.
XX 18-JUN-1999; 99US-0139454.
XX 18-JUN-1999; 99US-0139455.
XX 18-JUN-1999; 99US-0139456.
XX 18-JUN-1999; 99US-0139457.
XX 18-JUN-1999; 99US-0139458.
XX 18-JUN-1999; 99US-0139459.
XX 18-JUN-1999; 99US-0139460.
XX 18-JUN-1999; 99US-0139461.
XX 18-JUN-1999; 99US-0139462.
XX 18-JUN-1999; 99US-0139463.
XX 18-JUN-1999; 99US-0139750.
XX 18-JUN-1999; 99US-0139763.
XX 21-JUN-1999; 99US-0139817.
XX 22-JUN-1999; 99US-0139899.
XX 23-JUN-1999; 99US-0140353.
XX 23-JUN-1999; 99US-0140354.
XX 24-JUN-1999; 99US-0140695.
XX 28-JUN-1999; 99US-0140823.
XX 29-JUN-1999; 99US-0140991.
XX 30-JUN-1999; 99US-0141287.
XX 01-JUL-1999; 99US-0141842.
XX 01-JUL-1999; 99US-0142154.
XX 02-JUL-1999; 99US-0142055.
XX 06-JUL-1999; 99US-0142390.
XX 08-JUL-1999; 99US-0142803.
XX 09-JUL-1999; 99US-0142920.
XX 12-JUL-1999; 99US-0142977.
XX 13-JUL-1999; 99US-0143542.
XX 14-JUL-1999; 99US-0143624.
XX 15-JUL-1999; 99US-0144005.
XX 16-JUL-1999; 99US-0144085.
XX 16-JUL-1999; 99US-0144086.
XX 19-JUL-1999; 99US-0144325.
XX 19-JUL-1999; 99US-0144331.
XX 19-JUL-1999; 99US-0144332.
XX 19-JUL-1999; 99US-0144333.
XX 19-JUL-1999; 99US-0144334.
XX 19-JUL-1999; 99US-0144335.
XX 20-JUL-1999; 99US-0144352.
XX 20-JUL-1999; 99US-0144632.
XX 20-JUL-1999; 99US-0144884.
XX 21-JUL-1999; 99US-0144814.
XX 21-JUL-1999; 99US-0145086.
XX 21-JUL-1999; 99US-0145088.
XX 22-JUL-1999; 99US-0145085.
XX 22-JUL-1999; 99US-0145087.
XX 22-JUL-1999; 99US-0145089.
XX 22-JUL-1999; 99US-0145192.
XX 23-JUL-1999; 99US-0145145.
XX 23-JUL-1999; 99US-0145218.
XX 23-JUL-1999; 99US-0145224.
XX 26-JUL-1999; 99US-0145276.
XX 27-JUL-1999; 99US-0145913.
XX 27-JUL-1999; 99US-0145918.
XX 27-JUL-1999; 99US-0145919.
XX 28-JUL-1999; 99US-0145951.
XX 02-AUG-1999; 99US-0146386.
XX 02-AUG-1999; 99US-0146388.
```

```
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147192.
PR 07-AUG-1999; 99US-0147260.
PR 08-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 10-AUG-1999; 99US-0147493.
PR 11-AUG-1999; 99US-0147935.
PR 12-AUG-1999; 99US-0148171.
PR 13-AUG-1999; 99US-0148319.
PR 14-AUG-1999; 99US-0148341.
PR 15-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 19-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 21-AUG-1999; 99US-0149723.
PR 22-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 24-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151085.
PR 28-AUG-1999; 99US-0151086.
PR 29-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 02-SEP-1999; 99US-0152363.
PR 03-SEP-1999; 99US-0153070.
PR 04-SEP-1999; 99US-0153758.
PR 05-SEP-1999; 99US-0154018.
PR 06-SEP-1999; 99US-0154039.
PR 07-SEP-1999; 99US-0154779.
PR 08-SEP-1999; 99US-0155139.
PR 09-SEP-1999; 99US-0155486.
PR 10-SEP-1999; 99US-0155659.
PR 11-SEP-1999; 99US-0156458.
PR 12-SEP-1999; 99US-0156596.
PR 13-SEP-1999; 99US-0157117.
PR 14-SEP-1999; 99US-0158232.
PR 15-SEP-1999; 99US-0158369.
PR 16-SEP-1999; 99US-0159293.
PR 17-SEP-1999; 99US-0159294.
PR 18-SEP-1999; 99US-0159295.
PR 19-SEP-1999; 99US-0159329.
PR 20-SEP-1999; 99US-0159330.
PR 21-SEP-1999; 99US-0159331.
PR 22-SEP-1999; 99US-0159637.
PR 23-SEP-1999; 99US-0159638.
PR 24-SEP-1999; 99US-0159584.
PR 25-SEP-1999; 99US-0160741.
PR 26-SEP-1999; 99US-0160767.
PR 27-SEP-1999; 99US-0160768.
PR 28-SEP-1999; 99US-0160770.
PR 29-SEP-1999; 99US-0160815.
PR 30-SEP-1999; 99US-0160980.
PR 01-OCT-1999; 99US-0160981.
PR 02-OCT-1999; 99US-0160989.
PR 03-OCT-1999; 99US-0161404.
PR 04-OCT-1999; 99US-0161405.
PR 05-OCT-1999; 99US-0161406.
PR 06-OCT-1999; 99US-0161359.
PR 07-OCT-1999; 99US-0161360.
PR 08-OCT-1999; 99US-0161361.
PR 09-OCT-1999; 99US-0161920.

Query Match 19.5%; Score 359.8; DB 21; Length 1533;
Best Local Similarity 58.7%; Pred. No. 3.7e-100; Indels 12; Gaps 3;
Matches 682; Conservative 0; Mismatches 467;

Qy 676 GAACCTCTTTGTCGCGGAAACAAACCTTGACCACGCTCCAGAGTCCAGGCTGAAACTCTCCCT 735
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 319 GATCTTGTGTACCGAACCGAGAGCGGAGAGAAACACGACGCGGAGATTGGCG 378

Qy 736 TCATTATCAATTAAGCTGGATCTCCAGTGGGTCCAGGTTTTCAGGAGAGCTGGGCC 795
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 379 AGAGTGAGATTGACGCGGATTGATTGCAATGGATGCTGTTGAGTGAGTGGGCT 438

Qy 796 ACTCCCTCAAAGGTTTCATCGCGGAGAGAGGACTACAGGTTATGACCTTTGA---- 851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 439 AGTCCTCTTCGTGGGTTTATGAGGGAATCTGAGTTCCTCCAACTCTTCATTTCAATTTG 498

Qy 852 --CACCTCTAGATGATGGCGTGATCAACATAGCATCCCATTTGACTCCCGCTCTCT 909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 TTGAATCTCGATGATGGGTCTGTTGTTAATATGCTGTGCGCTATTGTTCTTGGCATTGAT 558

Qy 910 GCAGAGATAGACACGCGTGGAAAGGTGCAGCAAGTTTGTCTGGGCACATGGTGGACGG 969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 GATCAACAAAAGCCCTAATCGGTGAATCTAAACGTGTCTCCCTTGTGTTCTGATGAT 618

Qy 970 AG---GGTAGCTATCTTACGAGACGCTGAATTTCTATGAACACAGAAAGAGGACGCTGT 1026
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 AATCCAATCGTATCTCAATGATATTGAGATTTTAAACATCCGAAGAAGAGCGAATA 678

Qy 1027 TCCCGTCTTTGGGGACAACATGTACAAAACACCCCATATCAAAATGGTGAT---GGAA 1083
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 GCGAGAACTTGGGCTACGACTGCACCGGTTTGCTTATGTAGAGAGGCGATAACCAAT 738

Qy 1084 AGTGGGACTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAGAAAATAGATGAATGAT 1143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 GCTGGAGACTGGCTCATTTGGGGGTGATCTTGGAACTTTTGAACCTGTTAAGTACAATGAT 798

Qy 1144 GGGCTGGACCAATACCGCTCTGACACCTCTGAGCTCAACAGAAATGTAAGAAATGAAT 1203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 799 GGGCTTGATCGTTTCAGGCTTTCCCGCTTTGAACTGCGTGAAGAGCTAGAGAACTGGT 858

Qy 1204 GCTGATCGGTGTTTGCAATTCAGTTCGCAATCTGTCCACAATGCCATGGCTGTG 1263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 859 GCGGATCGGTCTTTGGCTTTACAGCTTAGGAACCCAGTTTCATATGACATGCTCTTCTT 918

Qy 1264 ATGACAGACACTGCCCGCAGGCTCTAGAGAGGGCTACAGCACCAGGCTCCTCTACTA 1323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 919 ATGACTGATATCTGTAGGAGACTTCTTGATGGGTTTATAAAACCTATCCTTTTGTCTT 978

Qy 1324 CAGCCTCTGGCGGCTGGACCAAGGATGACGATGTGCTCTAGACTGGCGGATGAAGCAG 1383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 979 CATCCATTGGGAGGTTTACAAAGCTGATGATGTTCTCTAGCTGGCGAATGAACAG 1038

Qy 1384 CACGGGCTGTGCTCGAGGAAGGGTCTCGATGCCCAAGTCAACCATTTGTCATCTTT 1443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1039 CAGGAGAAAGGTGTAGAGGATGGTGTCTTGATCCAGAGACTACTGTGGTTTTCATATTC 1098

Qy 1444 CCGTCTCCCATGTTATATGCTGGCCCCACAGAGTCCAGTGGCACTGCAGTCCCGGATG 1503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1099 CCAATCTCAATGCTCTATGCTGGTCCAAACGAGTCCAAATGGCAACGCAAGGTAGGATG 1158

Qy 1504 ATTGCGGGTGCCAAATTTCTACATTTGGGGAGGAGACCTCTCAGGAATGCCCATCTGAA 1563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1159 AATGCTGGTGTAACTTCTACATTTGTCGGTAGGATCCGGCTGGAATGGGTGATCCCGGTG 1218

Qy 1564 ACCAAGAAGGATCTGTATGAACCCATCATATGGGGGCAAGTCTTTGAGCATGGCCCTGGC 1623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1219 GAGAAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1278

Qy 1624 CTCACCTCTCTGGAAATCATTTCCATTCGAGTGGCTGCCTACAAACAGCAAAAGGCC 1683
```

Db	1279	CTCGAACGACTCAACATTCTTCCTTTCAGGGTGTGCTACGATAAGACACAAGGAAAG	1338
QY	1684	ATGGACTTCTATGATCCAGCAAGGCAACATGAGTTTGACTTCATCTCAGGAACTCGAATG	1743
Db	1339	ATGGCTTTCTTTGATCCCTCAAGGGCTCAGGACTTCTTGTTCATTTCTGGCACAAGATG	1398
QY	1744	AGGAAGCTCGCCGGGAAGGAGAGATCCCCAGATGGCTTCATGGCCCCCAAGCATGG	1803
Db	1399	AGAGCATTTGGCAAGAAACAGAGAACCCACCAGATGGATTATGTGCCCTGGAGGCTGG	1458
QY	1804	AAGTCCTCACAGATTATTAC	1824
Db	1459	AAAGTCCTTGTGGATTACTAC	1479

Search completed: June 4, 2003, 15:23:01
 Job time : 457 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Query Match 53.7%: Score 990; DB 1: Length 2506;
 Best Local Similarity 71.3%; Pred. No. 3.7e-305;
 Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;

QY	9	GATCAAGAGCAAAAGACGGAGAACACAGCAGAGAAATCCACCAATAGTCTCTATCAGGCCCA	68
Db	74	GAGCAATAACGCGCAGAACTGGGAATGCAGAGAGCAACCAATGTCACCTACCAAGCCCA	133
QY	69	CCATGTCAGCAGGAAATAGAGAGGGCAAGTGGTTGGACAAAGGGTGGTTCCAGGATG	128
Db	134	TCATGTCAGCAGGAAACAGAGAGTCAAGTGGTGGGACCAAGAGTGGCTTCCTGGTGG	193
QY	129	TACCGTGGCTAACAGTCTCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	188
Db	194	CACAGTTGGCTAACAGCTTCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	253
QY	189	GGAGTACCTGTCTCCATGCCATCCCTCTGTCTCTGTCTGTCTGTCTGTCTGTCTGTCT	248
Db	254	GGAGTACCTGTCTCCATGCCATCCCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	313
QY	249	TGGCCTTAACAGAAATCTGGATCTCTCTCTGGGACAGAGAGAAATATPCGCGGAT	308
Db	314	AGGTCTCAATTAATAATCTGGCTTTAGTCTCTGGGACAGAGAGAAATATPCGCGGAT	373
QY	309	TGCTGAGGTGGCTAAGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	368
Db	374	CGCAGAGTTGCTAAGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	433
QY	369	TCCATTCGCAAGAGATCGTGAGATGCCCGCAAAATACATGACAGGGGTGCCATT	428
Db	434	ACCTTACACTCAGGATCGCAACAAATCAAGGCAAAATTCATGAAGTGCAGTTACCGTT	493
QY	429	CTTTGAAATATTTAGATGCACTCTAATATTTTGTGAAGCAGAGACGCTAAAGCCCT	488
Db	494	TTTGTGAAGTATTTGATGCTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT	553
QY	489	CTATAAAGGGCCGAGCTGGGAGATTAAGGATTTACAGGTATTGATTTCTGATATGA	548
Db	554	CTACAAAAGCCGGCAGAGAAATTAAGGTTTCACTGGGATGATTTCTGAATATGA	613
QY	549	GAACTGAACTCTCTGAGCGTGTCTTAAACCAATTTGTCCACAGTGAAGTGTGT	608
Db	614	AAAGCCAGAGGCCCTGAGTTGGTGTCTGAAACAGAGCTCTCTGTGATTAATGACTGTGT	673
QY	609	CCACAGGTAGTGAAGTCTGCAAGCAGACATTTGACCTTATCTATATCAATCAAGA	668
Db	674	CCAGCAAGTTGTGGAAGTCTTACAGGAAGGGATATTTGACCTGTGGATGCACTTTATGA	733
QY	669	TATCCACGAACTCTTTGTCCGGAACAAACTTTGACACAGCTCCGAGCTGAGCTGAAAC	728
Db	734	AGTAAAGAACTATATGTCCCAAAATAAATCACTTTGCGCAAAACAGATCGGGAAC	793
QY	729	TCCTCCCTCATATCAATTAAGTGGATCTCCAGTGGTCCAGTGGTGGTGGTGGTGGTGGT	788
Db	794	ATTACCAAGCACTGAAATTAATAAGTGGATATGCAAGTGGTGGTGGTGGTGGTGGTGGT	853
QY	789	CTGGCCACTCCCTCAAGGTTTCATCGGGAGAGAGTACTTACAGTTATGACATT	848
Db	854	TTGGGCAACCCCATTAATGGCTTTATGAGAGAGGGAGTACTTCCAGTGGCTTCATT	913
QY	849	TGACACCCCTGCTAGATGGCTGTATCAACATGAGCATCCCATTTGTACTGCCGCTC	908
Db	914	TGATTTCTCTGGATGGAGGTGCTAATTAAGTGTGCTAGTACCTATAGTCTGACTCGAC	973
QY	909	TGCAGAGTAAGACAGCGCTGGAAGGGTGCACAAAGTTTGTCTGGCAGATGGTGGAGC	968
Db	974	TCATGAAGATAAAGAGAGCTGGACGGCTGTACAGCATTTGCTCTGATGATGAGGCCG	1033
QY	969	GAGGTAGTATCTTACGAGACCTGCAATTTCTATGACACAGAAAGAGGAGCGTGTTC	1028
Db	1034	CCGTGTGGCCATCTCTCGCAATCCAGAGTTTGTGAGCAGAGAAAGAGGAGCGGTGTC	1093
QY	1029	CCGTGTGGGGGACAAACATGTACAAACACCCCATATCAAAATGCTGATGAAAGTGG	1088

Db	1094	CAGACAGTGGGGAACGACATCAAGAACCACCCCTATATTAAAGTGGTGAACAAGG	1153
QY	1089	GGACTGCTGGTGGTGGAGACCTTCAGGTCTCGAGAAAATAAGATGGATGATGGCT	1148
Db	1154	AGATTGCTGATGGAGAGATCTCAAGTCTTGGATCGAGTTTATTGGAATGATGGTCT	1213
QY	1149	GGACCAATACCGTCTGACACCTCTGGAGCTCAACAGAAAATGTAAGAAAATGAATGCTGA	1208
Db	1214	TGATCAGTACGCTTACTCTCTACTAGCTAAGCAAGCAAAATTAAGATATGATGCTGA	1273
QY	1209	TGCGGTGTTTGCATTTCCAGTTGCCATCTCTGTCACAAATGGCCATGCCCTGTTGATGCA	1268
Db	1274	TGCTGTCTTTGCAATTCACACTACCAACCCAGTGCACAAATGGACATGCCCTGTTAATGCA	1333
QY	1269	GGACACCTCGCGAGGCTCTAGAGAGGGCTACAAAGCCCGCTCTCTCTACTTACACCC	1328
Db	1334	GGATACCCATAAGCAACTTCTAGAGAGGGCTACCGGGCCCTCTCTCTCTCTCTCTCTCT	1393
QY	1329	TCTGGGGGCTGGACCAAGGATGACGATGTCCCTCTAGACTGGCGGATGAAGACGACGC	1388
Db	1394	TCTGGGTGGCTGGCAAAAGGATGACGATGTCTCTTCTGATGTGGGTATGAAGCAGATGC	1453
QY	1389	GGCTGTGCTCGAGAGGGGTCTCTGGATCCCAAGTCAACCATTTGTTGCCATCTTCCGTC	1448
Db	1454	TGCAGTGTGGAGGAGGAGTTCTGAATCTGAGACGACAGTGGTGGCCATCTTCCCATC	1513
QY	1449	TCCCATGTTATATGCTGGGCCCCACAGAGTCCAGTGGCTGCAAGTCCCGGATGATTGC	1508
Db	1514	TCCCATGATGATGCTGGACCACTGAGTCCAGTGGCATTCAGACGACGATGTTGTC	1573
QY	1509	GGGTGCCAATTTCTACATTTGGGGAGGACCTTCAGGAATGCCCATCTTGAACCAA	1568
Db	1574	AGGAGCAACTTTTACATTTGAGCAGACCTGCTGGCATGCTTCATCCAGAAACAGG	1633
QY	1569	GAAGATCTGTATCAACCCACTCATGGGCAAGTCTTTCAGCATGCCCTGGCCTCAC	1628
Db	1634	GAAGGATCTTTATGAGCCAAAGTCTGTCGCAAGTCTGACGATGGCCCTGGTTAAT	1693
QY	1629	CTCTGTGGAATCATTTCCATTTCCGAGTGGCTGCTACAAACAAAGCCAAAAGCCATGGA	1688
Db	1694	CACCTTTGGAATAGTTCCCTTCTGAGTTCAGCTTACAAAGAAAGAGCGATGGA	1753
QY	1689	CTTCTATGATCCAGCAGGCAATGAGTTGACTTCATCTCAGGAATCGAATGAGGAA	1748
Db	1754	CTACTATGATCTCTGAACACCATGAAGACTTTGAATTTATTCAGAAACGAATGCGCA	1813
QY	1749	GCTCGCCGGGAGGAGAGAAATCCCCAGATGGCTTCATGGCCCCCAAGCATGGAGGT	1808
Db	1814	ACTTGTCTGAGAGGCGCAAAACCATGAAAGTTTTCATGGCTCCCAAGCTTGGACCGT	1873
QY	1809	CCTGACAGATTTATACAGTCCCTGGAGAA	1838
Db	1874	GCTGACAGATACTACAAATCCTTTGGAGAA	1903

RESULT 2
 US-09-346-408-5
 ; Sequence 5, Application US/09346408B
 ; Patent No. 6338966
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Steve
 ; APPLICANT: Anderson, Shawn
 ; APPLICANT: Rafalski, Antoni
 ; APPLICANT: Falco, Carl
 ; FILE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
 ; CURRENT APPLICATION NUMBER: US/09/346,408B
 ; EARLIER FILING DATE: 1999-07-01
 ; EARLIER FILING DATE: 60/092,833
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: Microsoft Office 97

; SEQ ID NO 5
; LENGTH: 1697
; TYPE: DNA
; ORGANISM: Glycine max
US-09-346-408-5

Query Match 20.8%; Score 383; DB 4; Length 1697;

Best Local Similarity 59.9%; Pred. No. 9.8e-112;

Matches 698; Conservative 0; Mismatches 455; Indels 12; Gaps 3;

```
QY 671 TCCAGAACTCTTTGCGGAAACAAACAACTTGACCACGTCGAGCTGAGGCTGAACATC 730
D 671 TCCAGAACTCTTTGCGGAAACAAACAACTTGACCACGTCGAGCTGAGGCTGAACATC 730
D 315 TCGTGAGCTTGTGTGAGGATTTGAGAGGATTTGAAGAAGGCTGAGGCTCTTTCTGT 374
QY 731 TCCCTTCTATATCAATTAATCACTAGCTGGATCTCCAGTGGCTCAGGTTTGTAGCGAAGGCT 790
D 731 TCCCTTCTATATCAATTAATCACTAGCTGGATCTCCAGTGGCTCAGGTTTGTAGCGAAGGCT 790
D 375 TCCCGAGGATCAAGCTCTCAAGGATGACCTTGAGTGGTCCATGCTCTCCAGCAAGGAT 434
QY 791 GGGCCACTCCCTCAAGGTTTTCATGCGGGAGAGGATGACATGACAGTATGACAGTATG 850
D 791 GGGCCACTCCCTCAAGGTTTTCATGCGGGAGAGGATGACATGACAGTATGACAGTATG 850
D 435 GGGCCACACCCCTGAAAGGCTTCATGAGAGAGCGAGTTCCTCCAAACGCTTCATTTCA 494
QY 851 ACACCTCGTA-----GATGATGCGTGAATCAACATGAGCATGACATGACATGACATG 904
D 851 ACACCTCGTA-----GATGATGCGTGAATCAACATGAGCATGACATGACATGACATG 904
D 495 ACTCGCTCGATCGATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 554
QY 905 TCTCTGAGAGGATGAGACACGCTGGAAGGTCGAGCAAGTTTGTCTGGCACATGG-- 962
D 905 TCTCTGAGAGGATGAGACACGCTGGAAGGTCGAGCAAGTTTGTCTGGCACATGG-- 962
D 555 TTGATGATGCGCAGAGCATCGATCGGGGATAACAAAGGTTCTCTTTTGATTCCA 614
QY 963 -TGGAGGAGGAGTACTATCTTACGAGAGCTGTAATTTATGAACACAGAAAGAGAAC 1021
D 963 -TGGAGGAGGAGTACTATCTTACGAGAGCTGTAATTTATGAACACAGAAAGAGAAC 1021
D 615 AGGAGACCCCGTTGCAATTCATTAATATGAGATTATAAGCATCTTAAGAGAGAA 674
QY 1022 GCTGTTCCGCTTTGGGGCAACATGTACAAACACCCCATATC---AAATGGTCA 1078
D 1022 GCTGTTCCGCTTTGGGGCAACATGTACAAACACCCCATATC---AAATGGTCA 1078
D 675 GAATAGCCGGAATTTGGGGAACCATTTGCCCTGAGCTTATGTTGAACAAATATAA 734
QY 1079 TGGAAAGTGGGACTGGCTGGTGGTGGAGACTTTCAGTGTGGAGAAATAAGATGGA 1138
D 1079 TGGAAAGTGGGACTGGCTGGTGGTGGAGACTTTCAGTGTGGAGAAATAAGATGGA 1138
D 735 CCAATGCTGGAATTTGTTGATTTGGGGTACCTAGAGTCAATTAACCAATTCAGTACA 794
QY 1139 ATGATGGGTGACCAATACCGTCTGACACCTCTGAGCTCTGAGCTCAACAGAAATGAAGAA 1198
D 1139 ATGATGGGTGACCAATACCGTCTGACACCTCTGAGCTCTGAGCTCAACAGAAATGAAGAA 1198
D 795 ATGATGGGATGATCATTTTCTCTATCTCGGCACAACTCGTGCAGAGTTCAAGGC 854
QY 1199 TGAATGCTGATGCGGTGTTGATTCAGTTCGCAATCTGTCACATGTCACATGSCCATGCC 1258
D 1199 TGAATGCTGATGCGGTGTTGATTCAGTTCGCAATCTGTCACATGTCACATGSCCATGCC 1258
D 855 GCAATGCGGATGCTGTTTGGCTTCCAGCTCCGGAATCTGTTTCAATGGCCATGCTT 914
QY 1259 TGTGATGACGACACCTGCGCGAGGCTCTAGAGAGGCTTACAGACACCGGTCCTCC 1318
D 1259 TGTGATGACGACACCTGCGCGAGGCTCTAGAGAGGCTTACAGACACCGGTCCTCC 1318
D 915 TGTATGATGACGACACCGGAAAGCGCTTCTGAGATGGCTTAAAGATCTGTCCTCT 974
QY 1319 TACTACACCTCTGGGGCTGGACCAAGGATGAGATGCTCTAGCTAGCTGGCGGATGA 1378
D 1319 TACTACACCTCTGGGGCTGGACCAAGGATGAGATGCTCTAGCTAGCTGGCGGATGA 1378
D 975 TGTCTTATCCACTTGGAGGCTACACCAAGCTGATGATGCCACTTGTATGGCAATGA 1034
QY 1379 AGCAGCAGCGGCTGCTGTCGAGGAGGCTCTGATGCCCAAGTCAACCATTTGTGCA 1438
D 1379 AGCAGCAGCGGCTGCTGTCGAGGAGGCTCTGATGCCCAAGTCAACCATTTGTGCA 1438
D 1035 AGCAACATGAGAGGATCTGAGGATGTTGTTCTGATCCAGAGAACTGTGGTATCCA 1094
QY 1439 TCTTCCGCTCTCCATGTTATATGCTGGCCACAGAGGCTCCAGTGGCACTGAGGTC 1498
D 1439 TCTTCCGCTCTCCATGTTATATGCTGGCCACAGAGGCTCCAGTGGCACTGAGGTC 1498
D 1095 TATTTCCATCTCCATGACATGCTGGACCCACGAGGTGTCAGTGGCATGCAAGGCTA 1154
QY 1499 GGATGATGCGGGTCCCAATTTCTATCTTGGGGAGGACCCCTGCAGGAATGCCCATPC 1558
D 1499 GGATGATGCGGGTCCCAATTTCTATCTTGGGGAGGACCCCTGCAGGAATGCCCATPC 1558
D 1155 GGATCAATGCGGGCTAACTTCTATATGCTGTTGGTGGTGGTGGTGGTGGTGGTGGT 1214
QY 1559 CTGAACCAAGAGGATCTGTATGAACCCACTCATGGGGGCAAGCTCTTGACATGGCC 1618
D 1559 CTGAACCAAGAGGATCTGTATGAACCCACTCATGGGGGCAAGCTCTTGACATGGCC 1618
D 1215 CAGTTGAGAAAGAGATCTGTATGATGCTGACCATGGAAAGAAAGTATTTGAGCATGGC 1274
```

RESULT 3

US-09-346-408-7

; Sequence 7; Application US/09346408B

; Patent No. 638966

; GENERAL INFORMATION:

; APPLICANT: Allen, Steve

; APPLICANT: Anderson, Shawn

; APPLICANT: Falco, Carl

; APPLICANT: Rafalski, Antoni

; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins

; FILE REFERENCE: BB-1167-A

; CURRENT APPLICATION NUMBER: US/09/346,408B

; CURRENT FILING DATE: 1999-07-01

; EARLIER APPLICATION NUMBER: 60/092,833

; EARLIER FILING DATE: July 14, 1998

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 7

; LENGTH: 1890

; TYPE: DNA

; ORGANISM: Triticum aestivum

US-09-346-408-7

Query Match 20.0%; Score 368.2; DB 4; Length 1890;

Best Local Similarity 58.9%; Pred. No. 5.5e-107;

Matches 694; Conservative 0; Mismatches 473; Indels 12; Gaps 3;

```
QY 671 TCCAGAACTCTTTGCGGAAACAAACAACTTGACCACGTCGAGCTGAGGCTGAACATC 730
D 671 TCCAGAACTCTTTGCGGAAACAAACAACTTGACCACGTCGAGCTGAGGCTGAACATC 730
D 170 TCGTCGACCTCGTGGCGCGCGCGGAGCGCGCGCGCTCGCGCGCGCGCGCGCGCG 229
QY 731 TCCCTTCAATTAATTAAGCTGGATCTCCAGTGGGTCCAGGTTTGTAGCGAAGGCT 790
D 731 TCCCTTCAATTAATTAAGCTGGATCTCCAGTGGGTCCAGGTTTGTAGCGAAGGCT 790
D 230 TCCCGCGGTGCGGCTCGCGCGCGGTGGAGTGGGCGCACGTGCTCGCGGAGGCT 289
QY 791 GGGCAGCTCCCTCAAGGTTTTCATGCGGGAGAGGATGACTTACAGGTTATGCACTTTG 850
D 791 GGGCAGCTCCCTCAAGGTTTTCATGCGGGAGAGGATGACTTACAGGTTATGCACTTTG 850
D 290 GGGCGTCCCGCTGCGCGGCTTCATGCGGAGCAGAGTACCTCAGTGCCTCCACTTCA 349
QY 851 ACAC-----CCTGCTAGATGATGCGTCAACATGAGCATGCCATTTGACTGCCCG 904
D 851 ACAC-----CCTGCTAGATGATGCGTCAACATGAGCATGCCATTTGACTGCCCG 904
D 350 ACTCCCTCGGCTCTCCCTCCGCGGCTCGCCAAACATGCTGCTGCCCATCTGCTCGCG 409
QY 905 TCTCTGAGAGGATAAGACACGGCTGGAGGCTGACGCAAGTTTGTCTTGGC---ACATG 961
D 905 TCTCTGAGAGGATAAGACACGGCTGGAGGCTGACGCAAGTTTGTCTTGGC---ACATG 961
D 410 TCGACGACGCGCCCAAGGACCGCTCGCGCGCGCGCGCGCGCTCGCGCGCGCGCG 469
QY 962 GTGAGGAGGAGTACTATCTTACAGAGCTGTAATTTCTATGAACACAGAAAGAGAAC 1021
D 962 GTGAGGAGGAGTACTATCTTACAGAGCTGTAATTTCTATGAACACAGAAAGAGAAC 1021
D 470 AGGCGAGCTCTCTCGCGCTTCCTCCGAGTGTGCAATATACCTTCACATAAAGAGAA 529
QY 1022 GCTGTTCCGCTTTGGGGCAACATGTACAAACACCCCATATCAAAATGGTGAT-- 1079
D 1022 GCTGTTCCGCTTTGGGGCAACATGTACAAACACCCCATATCAAAATGGTGAT-- 1079
D 530 GGATGCAAGAAACATGGGGGCAACACTGGCGCTTACCTTATGTCATGAGCGCATAA 589
```


Db 1262 GATAAACTCAGGCAAAATGGCTTTTCGATCCCTCGAGACCTCAGGATTTCTCTTTTC 1321
QY 1726 ATCTCAGCAACTCGAATCAGGAAGCTCCGCCGGAAGGAGAGAAATCCCCCAGATGGCTTC 1785
Db 1322 ATATCCGCAACCAAGATGAGACACTACCAAGAACAAGAGATCCACCAGAGGGTTT 1381
QY 1786 ATGCGCCCAAGCATGGAAGTCTCTGACAGATTATTA 1823
Db 1382 ATGTGCCCCAGGGATGCAAGGTGTGTGTTGAATATTA 1419

RESULT 5

US-09-149-476-24
Sequence 24, Application US/09149476
Patent No. 6420526
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 186 Human Secreted proteins
FILE REFERENCE: P2002PI
CURRENT APPLICATION NUMBER: US/09/149,476
CURRENT FILING DATE: 1998-09-08
EARLIER APPLICATION NUMBER: PCT/US98/04493
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,615
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,892
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/047,595
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,669
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/049,610
EARLIER FILING DATE: 1997-06-13
EARLIER APPLICATION NUMBER: 60/061,060
EARLIER FILING DATE: 1997-10-02

Query Match 12.2%; Score 226; DB 4; Length 2323;
Best Local Similarity 92.2%; Pred. No. 1.4e-61;
Matches 271; Conservative 0; Mismatches 20; Indels 3; Gaps 3;

QY 1 ATGTGGGGATCAAGAAGCAAAAGCGAGAACCGAGAGAAATCCACCAATGTAGTCTAT 60

Db 165 ATGTGGGGATCAAGAAGCAAAAGCGAGAACCGAGAGAAATCCACCAATGTAGTCTAT 224
QY 61 CAGGCCACCATGTGACGAGGAATAAGAGAGGGAAGTGGTTGGAACAAGGGTGGGTTTC 120
Db 225 CAGGCCACCATGTGACGAGGAATAAGAGAGGGAAGTGGTTGGAACAAGGGTGGGTTTC 284
QY 121 CGAGGATGTACCGTGTGGCTTAACAGGTCTCTCTGGTGTCT -GGAAAAACAACGATAAGTTT 179
Db 285 CGAGGATGTACCGTGTGGCTTAACAGGTCTCTCTGGTGTGGGAAAAACAACGATAAGTTT 344
QY 180 TGCCTCGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTCGATGGGAGCAA 239
Db 345 TGCCTCGAGGAGTA-CTTGTCTCCCATGCCATCCC-TGTTAATTCCTGGATGGGAGCAA 402
QY 240 TGTCCTGATGGCTTAACAGAAATCTCGGATCTCTCTGGGACAGAGGA 293
Db 403 TGTCCTGATGGCTTAACAGAAATCCCCAGATGGCTTCATGGCCCCCAAGCA 456

RESULT 6

US-09-346-408-1
; Sequence 1, Application US/09346408B
; Patent No. 638966
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falco, Carl
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-A
; CURRENT APPLICATION NUMBER: US/09/346,408B
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 60/092,833
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (479)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (495)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (508)
US-09-346-408-1

Query Match 10.7%; Score 197.4; DB 4; Length 516;
Best Local Similarity 65.6%; Pred. No. 7e-53;
Matches 320; Conservative 0; Mismatches 162; Indels 6; Gaps 2;

QY 1140 TGATGGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAGAGAAATGTAAGAAAT 1199
Db 1 TGATGGCTAGATCAGTATCGCTCTCCAGCACAGCTCGGTGAAGAGTTTGCCAGCG 60
QY 1200 GAATGCTGATGCGGTGTTTGCATTCAGTTGCGCAATCCCTGTCCACATGGCCATGCCCT 1259
Db 61 CAATGCTGATGCAATTTTGCCTTCAGCTTCGCAATCCAGTACACAAATGGCATGCTCT 120
QY 1260 GTTGATGACGAGCACCTGCGGAGGCTCTAGAGAGGGGTACAAAGCACCCGGTCCCTCT 1319
Db 121 ACTATGACGAGCACAGGCAACGCCCTCTTGAGATGGGTTTATAAACCCTGTTCTCT 180
QY 1320 ACTACACCTCTGGGCGGCTGGACCAAGGATGACGATGCGCTCTAGACTGGCGGATGAA 1379
Db 181 GCTCCATCCACTGGGAGGATTCACAAAGCAGATGATGTCCTCTTAGTTGAGAAATGAA 240
QY 1380 GCAGCAGCGGCTGTCTCGAGGAAGGGTCTGGATGCCCAAGTCACACCATTTGTTGCCAT 1439

Db 241 GCAACATGAGAAGTCTTCTGAGGAAGGTGCTCAACCCAGAAATCAACTGTTCTGGCAT 300
QY 1440 CTTTCCTGTCCTCCATGTTATATGCTGGCCCCACAGAGTCCAGTGGCACTGCAAGTCCCG 1499
Db 301 CTTCCCTCCTCATGATGCTGGCCCACTGAGTTCAATGGCATCTAAGGCTCG 360
QY 1500 GATGATGGGGTGCCAAATTCATATGTTGGGAGGAGCCCTGCAGGAATGCCCATCC 1559
Db 361 TATTAATGCTGGCGCAAAATTCATATGTTGGAAGGATCTGCTGTATGAGCACCCAC 420
QY 1560 TGAACCAAGAAGATCTGATGAACCCACTCATGGGGGCAAGTCTTGAGCATGGCCC 1619
Db 421 AGAGA-----AAGAGGACTCATGCTGATCATCGGAAAGAAGGT-TTGAGCATGGCTCC 474
QY 1620 TGGCTCA 1627
Db 475 TGCNTGA 482

RESULT 7

US-09-153-310-41
; Sequence 41, Application US/09153310
; Patent No. 6326184
; GENERAL INFORMATION:
; APPLICANT: Gjermansen, Claes
; Hansen, Jorgen
; Johannesen, Pia Francke
; Pedersen, Mogens Bohl
; Sorensen, Steen Bech
; TITLE OF INVENTION: Method of producing a composite
; fermented beverage using genetically modified yeast
; strains
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/153,310
; FILING DATE: 15-Sep-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-672-5300
; TELEFAX: 202-672-5399
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1160 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-153-310-41

Query Match 9.2%; Score 169.2; DB 4; Length 1160;
Best Local Similarity 57.6%; Pred. No. 1.1e-43;

Matches 303; Conservative 0; Mismatches 223; Indels 0; Gaps 0;
QY 124 GGATGTACCGTGGGTAAACAGGTCTCTCTGGTCTGGAAAAACAACGATAAGTTTGGC 183
Db 447 GGTGTACCGTGGGTAAACAGGTCTCTCTGGTCTGGAAAAACAACGATAAGTTTGGC 506
QY 184 CTGAGGAGTACCTGTTCTCCCATGCCATCCCTTGTGTTACCTCCCTGGAGGCAATGTC 243
Db 507 CTGGAACAATTAAGTCTCAAAAAAAGTATCTGCTTATAGGTTAGATGGTGAACATT 566
QY 244 CGTCATGGCCCTTAACAGAAATCTCGGATTCCTCTCTGGGACAGAGAGGAAAAATATCCGC 303
Db 567 CGTTTGGTTGAATAAGGATTTGGGCTTCTCAGAAAAAGGACAGAAATGAAAAACATTCGT 626
QY 304 CGGATTCCTGAGGTGGCTAAAGCTGTTTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 363
Db 627 AGAATTAGTCAAGTATCAAGCTATTCGCTGATTCGCTGATTCGCTGATTCATCCATTCATT 686
QY 364 ATTTCTCCATTCGAAAGGATCGTGAGAAATGCGCGGCAAAATACATGAATCAGCAGGCTG 423
Db 687 ATTTCCCATACAGAGTCGATAGACAGACAGCCCGTGTATTTACATAAGGAGCAGGCTTG 746
QY 424 CCATTCTTTCGAATATTTGTAGATGCACCTCTAAATATTTGTGAAACGACAGACGTAATA 483
Db 747 AGTTTCATGAAATTTTTTTTGTGATGTTTCCATTTAGAGTCGCTGAGCAAGAGACCCCTAAG 806
QY 484 GGCCTCTATAAAAGGGCGAGAGCTGGGAGATTAAGGATTTACAGGTATTTGATTCGTAT 543
Db 807 GGTTGTATTAAGAACCCAGAGAGGTGTGATTAAGAGTTCAGTGTATTTACAGTCTCT 866
QY 544 TATGAGAACTGAAACTCTGAGCGTGTGCTTAAACCAATTTGTCCACAGTGAGTGAC 603
Db 867 TAGCAAGCTCCAAAGGCCAGAGTTCATTTAAGAACTGACCAAAAGACTGTTTGAAGAA 926
QY 604 TGTGTCACCAAGGTAGTGGAACTTCTCAGAGCAGACACATTTGTAC 649
Db 927 TGTGCTGCTATCATTTATGAGTACCTGGTCAATGAGAAAGATTATCC 972

RESULT 8

US-09-346-408-9
; Sequence 9, Application US/09346408B
; Patent No. 6338966
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Anderson, Shawn
; APPLICANT: Falco, Carl
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Genes Encoding Sulfate Assimilation Proteins
; FILE REFERENCE: BB-1167-A
; CURRENT APPLICATION NUMBER: US/09/346.408B
; EARLIER FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 60/092,833
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 9
; LENGTH: 563
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (337)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (404)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (437)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (475)
; FEATURE:

Query Match 6.7%; Score 123; DB 4; Length 615;
Best Local Similarity 54.3%; Pred. No. 4.1e-29;
Matches 274; Conservative 0; Mismatches 225; Indel's 6; Gaps 1;

Db 1440253 GACTGT

1440253 GACTGTCCGGCTCCGGCAAGTCGTGGTGGCCATGCTGGTGGCCGAAGCTACTCGAA 1440

QY	206	ATGCCATCCCTGTTTACTCCCTGGATGGGACAATGTCGTCATGCGCTTACAGAAATC	265
Db	1440313	AGGGCATCTCCGCTTACGTTCTGGACGGGGAACAACCTAGCGATGCGCTCAACGGCGAC	1440372
QY	266	TCGGATTCTCTCCTGGGGACAGAGAGAAATATCGCGGGATTGCTGAGGTGGCTAAGC	325
Db	1440373	TGGGCTTTTCCATGCGCGACCGCGGAGAACCTGCGCGCGCTGTCGATGTGGCCACAC	1440432
QY	326	TGTTTGCTGATGTGCTGTGCTGCATTACCAAGCTTTATTCTCCATTGCGAAAGGATC	385
Db	1440433	TGCTCGCGGATTCTGCGCACCTGGTCTGTCGCCCGCATCAGCCCCCTGCTGAGCAC	1440492
QY	386	GTGAGAATGCCGCCAAATACATGAATCAGAGGGTGCCTATCTTTGAAATATTGTGATG	445
Db	1440493	GTCCCTTGCGTCGTAAGTGCACGCTGATCGGGAACTACATTTTTCGAGGTGTTCTGTG	1440552
QY	446	ATGCACCTCTAAATATTTGTGAAGCAGACAGCTAAAGGCCCTCTATAAAGGCCAGAG	505
Db	1440553	ACACCCGCTGCAGGACTGTGAGAGCGGTGATCCCAAGGGTGTACGCCAAAGCGGTG	1440612
QY	506	CTGGGGAGATTAAAGGATTTACAGTATTGATTCTGATTAGAAACCTGAAACTCCCTG	565
Db	1440613	CGGGTGAGATCAGCACTTCACCGGGATCGACAGCCCATATCAGCGGCCCAAGACCCAG	1440672
QY	566	AGCGTGTGTTAAACCAATTTGTCACAG	595
Db	1440673	ACCTACGGGTTACGCCGGATCGCAGCATAG	1440702

```

RESULT 11
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-00007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

Query Match	5.9%	Score 108.4	DB 4	Length 4403765
Best Local Similarity	50.8%	Pred. No. 4.9e-22		
Matches 259	Conservative 0	Mismatches 251	Indels 0	Gaps 22
QY	86	AGAGAGGGCAAGTGGTTGGAAACAAGGGGTGGGTTCGAGAGATGTACCGTGTGCTAACAG	145	
Db				
1439662	ACAGATCGTCTGCTCACTGCGCAAGATCGCGCGCCAGGGGCAAGCGGTGTGTTTACCG	1439721		
QY	146	GTCTCTCTGGTGTGTAAGAAACAACGATAAGTTTTCGCCCTGGAGGAGTACCTGTGCTCCC	205	
Db				
1439722	GACTGTCCGGCTCCGCAAGTCGTCTGGTGGCCATGCTGGTTGAGCGGAAGCTACTCGAA	1439781		
QY	206	ATGCCATCCCTTGTCTACTCCCTGGATGGGGACAAGTCCGTCATGCGCTTACAGAAATC	265	
Db				
1439782	AGGGCATCTCCGCTTACGTTCTTGAGCGGGACAACCTACGGGATGCGCTTCAACGGCGACC	1439841		
QY	266	TGCGATTCTCTCTCTGGGGACAGAGAGGAAATATCCGCGGGATTGCTGAGGTGGGTAAGC	325	

Db	1439842	TGGGCTTTTCCATTGGCCGACCGCGCGGAGAACTGCGCGGCTGTCGATGTGGCCACAC	1439901
QY	326	TGTTTGCTGATGCTGGTCTGGTCTGCATTACCAAGCTTTATTTCTCCATTCGCAAGGATC	385
Db	1439902	TGCTCGCGGANTGTGCCACCTGGTGTGGTCCGGCATCAGCCCCCTTGCTGAGCACC	1439961
QY	386	GTGAGAAATGCCCAAAATACATGATCAGCAGGGTGCCATTTCTTGAATATTGTTAG	445
Db	1439962	GTCCCTGGCTCGTAAGTGCACGCTGATCGGGAATCGACTTTTTCGAGGTTGTTCTGTG	1440021
QY	446	ATGCACCTCTAAATATTTGTAAAGCAGAGACGTAAAGGCCCTCTATAAAGGGCCAGAG	505
Db	1440022	ACACCCGGTGCAGGACTGTGAGAGCGGTATCCCAAAGGTTGTACGCCAAAGCGGTG	1440081
QY	506	CTGGGAGATTAAAGGATTTACAGGTATTCATCTGATTATGAGAACTGCAACTCCTTG	565
Db	1440082	CGGGTGAGATCAGCACTTCACCGGGATGACAGCCCATATCAGCGGCCCAAGACCCAG	1440141
QY	566	AGCGTGTGTTAAACCAATTTGTCACAG	595
Db	1440142	ACCTACGGGTACGCGGATCGCAGCATAG	1440171

RESULT 12
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 2.58; Score 46.6; DB 1; Length 7218;

THIS PAGE BLANK (user)

Query Match	98.3%	Score 1813.6;	DB 9;	Length 2617;
Best Local Similarity	59.0%;	Pred. No. 0;		
Matches 1841;	Conservative	0;	Mismatches 4;	Indels 15; Gaps 1
1	ATGTCGGGGATCAAGAAAGCAAAAGACGGAGAACCCAGACAGAAATCCACCAATGTAGTCTAT	60		
93	ATGTCGGGGATCAAGAAAGCAAAAGACGGAGAACCCAGACAGAAATCCACCAATGTAGTCTAT	152		
61	CAGCCCAACCATGTGAGCAGGAATAGAGAGGGCAAGTGGTTGGAAACAAGGGGTGGGTTTC	120		
153	CAGCCCAACCATGTGAGCAGGAATAGAGAGGGCAAGTGGTTGGAAACAAGGGGTGGGTTTC	212		
121	CGAGSAGTGACCGTGTGGCTAACAGGTCCTCTGTGGTGGAAAAACAACGATAAGTTTT	180		
213	CGAGSAGTGACCGTGTGGCTAACAGGTCCTCTGTGGTGGAAAAACAACGATAAGTTTT	272		
181	GCCCTGGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGGACAAT	240		
273	GCCCTGGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGGACAAT	332		

	Best Local Similarity	71.3%; Pred. No. 0;	Mismatches	Conservative	0;	Indels	525;	Gaps	0;
Qy	9	GATCAAGAACAAAGACCGGAGACCAGCAGAGAATCCACCNAATGTAGTCTATCATGAGGCCCA	68						
Db	94	GAGCAATAACCGCAGAACTGGGGAATGCAGAGAGCAACCAATGTCACCTACCAAGGCCA	153						
Qy	69	CCATGTGAGCAGGAATAAGAGGGCAAGTGTTTGGAAACAAGGGTGGTTCGCGAGGATG	128						
Db	154	TCAITGTCAGCAGGAACAGAGAGGTGAGTGGTGGGACCAGAGGTGGCTTTCGTGGTTG	213						
Qy	129	TACCGTGTGGCTTAACAGGTCTCTCTGTGTGTGGAANAACAACGATAAGTTTTGCCCTGGA	188						
Db	214	CACAGTTTGGCTTAACAGGCTTGTCTGGAGCGGGAAGACTACTGTGAGCATGGCCTTGGA	273						
Qy	189	GGAGTACCTTGTCTCCCATGCCATCCCTGTTACTCCCTGGATGGGACANATGTCGGTCA	248						
Db	274	GGAGTACCTGGTTGTCTATGGATTCCATGCTACACTCTGGATGGTGACAATATTCGTCA	333						
Qy	249	TGGCCTTTAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGAAAATATCCGCCGGAT	308						
Db	334	AGGTCTCAATAAAAAATCTTGCCTTAGTCTGAAGACAGAGAAGATGTTCCAGCGAT	393						
Qy	309	TGCTGAGGTGGCTAAGCTGTTTGGTGANCTGGTCTGTCTGCAATTACCAGCTTTATTTC	368						
Db	394	CGCAGAAAGTTGCTAAACTGTTTGAGATGCTGGCTTAGTGTGCATCACAACTTTCATATC	453						
Qy	369	TCCATTTCGCAAGAGATCGTGAGATGCCCGCAAATACATCAATCAGCAGGGCTTGCCATT	428						
Db	454	ACCTTACATCAGGATCGCAACAATGAAGGCAAAATTCATGAAGGTGCAAGTTTACCGCTT	513						
Qy	429	CTTTTGAATATTTGTAGATGCACCTCTAAATATTTGTGAAACGACAGACGTAAGAGGCT	488						
Db	514	TTTTTGAGTATTTGTTGATGCTCCTCTGTCATGTTTGTGAACAGAGGATGTCAAGGACT	573						
Qy	489	CTATAAAGGGCCAGAGCTGGGGAGATTTAAAGGATTTTACAGGTATTGATCTGTGATTATGA	548						
Db	574	CTACAAAAAGCCCCGGCAGAGAAATTAAGGTTTTCACCTGGGATCGATCTCGAATATGA	633						

QY 1629 CTCTGCGAATCATTCATCCGAGTGGCTGCTTACACAAAGCCAAAGCCATCGA 1688
Db 1714 CACTTGGAAATAGTTCCTTTCGAGTTGCGAGTTTACACAAAGAAAGACGCTATGGA 1773
QY 1689 CTTCTATGATCAGCAAGGCAATAGTTTTCGACTTTCATCTCAGGAATCGAATCAGGAA 1748
Db 1774 CTACTATGACTGTAACACCATGAGACTTTGAATTTATTCAGGAACACGATCCGCA 1833
QY 1749 GCTCGCCCGGGAAGAGAGATCCCCAGATGGCTTCATGCGCCCAAGGATGGAAGT 1808
Db 1834 ACTTGCTCGAAGGCGGAGAACACCATCTGAAGGTTTCATGGCTCCCAAGGCTTGACCGT 1893
QY 1809 CCGACATATTATTCAGTCCCTGGAGAA 1838
Db 1894 GCTGACAGATTAACAATCTTTGGAGAA 1923

RESULT 3

US-09-984-245-116
; Sequence 116, Application US/09984245
; Patent No. US20020165374A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE REFERENCE: P2004P1
; CURRENT APPLICATION NUMBER: US/09/984,245
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,350
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,188
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,135
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/050,937
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,187
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,099
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,352
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,186
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,069
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,095
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,131
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,096
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,355
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,160
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,351
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: US 60/048,154
; PRIOR FILING DATE: 1997-05-30

; PRIOR APPLICATION NUMBER: US 60/054,804
; PRIOR FILING DATE: 1997-08-05
; PRIOR APPLICATION NUMBER: US 60/056,370
; PRIOR FILING DATE: 1997-08-19
; PRIOR APPLICATION NUMBER: US 60/060,862
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 343
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 116
; LENGTH: 2527
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-984-245-116

Query Match 53.7%; Score 990; DB 9; Length 2527;
Best Local Similarity 71.3%; Pred. No. 0;
Matches 1305; Conservative 0; Mismatches 525; Indels 0; Gaps 0;
QY 9 GATCAAGAAGCAAAAGACGGAGAACACGACAGAAATCCACCATATGTAGTCTATCAGGCCCA 68
Db 94 GAGCAATAACCGCGCAGAACTGGGAATGCAGAGACCAACCAATGTCACTACCAAGCCCA 153
QY 69 CCATGTGAGCAGGAATAAGACAGGCGCAAGTGTGGAAACAAGGGTGGTCCGAGGATG 128
Db 154 TCATGTCAGCAGGAACAAGAGAGGTGAGTGGTGGGACAGAGGTGGCTTTCTGGGTG 213
QY 129 TACCGTGTGGCTAACAGGTCTCTCTGTGCTGGAAACAAACAGATAAGTTTTGCCCTGGA 188
Db 214 CACAGTTTGGCTAACAGGCTTGTCTGGAGCGGAAAGACTACTGTGACATGGCCTTGA 273
QY 189 GGAGTACTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAATGTCGCTCA 248
Db 274 GGAGTACTTGTGTCTCATGTATTCATCTGATCTGATGCTGATGCTGATGCTGATGCTG 333
QY 249 TGGCTTTAACAGAAATCTCGGATTTCTCTCTGGGACAGAGAGAAATATCCGCCGAT 308
Db 334 AGGTCTCAATAAATCTTGGCTTTAGTCTTGAAGACAGAGAGAAATGTTCCAGCGAT 393
QY 309 TGCTGAGTGGCTAAGCTGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 368
Db 394 CGCAGAAAGTTGCTAACTGTTTGCAGATGCTGGCTTAGTGTGCTGCTGCTGCTGCTGCTGCT 453
QY 369 TCCATTTCGAAAGGATCGTGAGATGCCCAAAATACATCAATCAGCAGGCTGCCATT 428
Db 454 ACCTTACACTCAGGATCGCAAAATCAAGGCAAAATTCATGAAGGTGCAAGTTTACCGTT 513
QY 429 CTTTGAATATTTGTAGATGACCTCTAAATATTTGTGAAGACAGAGACGTTAAAGGCT 488
Db 514 TTTTGAAGTATTTGTTGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 573
QY 489 CTATAAAGGCGCAGAGCTGGGAGATTAAGGATTTACAGGATTTACAGGATTTGATTTGATTA 548
Db 574 CTACAAAAAGCCCGCAGGAGAAATTAAGGTTTCACTGGGATCGATTTCTGAATATGA 633
QY 549 GAAACCTGAAACTCTGAGCTGCTGCTTAAACCAATTTGTCACAGCTGAGTGTGCTGT 608
Db 634 AAAGCCAGAGGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 693
QY 609 CCACCAAGTGTGGAATCTTCTGAAGACAGAGAACTTGTACCCCTATCTACTATAATCAAGA 668
Db 694 CCAGCAAGTTGTGGAATCTTCTACAGGAACGGGATATTTGACCTGTGGATGCATCTTATGA 753
QY 669 TATCCAGCAACTCTTTGTGCGGAAACAACTGTCACAGCTGCGAGCTGAGGCTGAAC 728
Db 754 AGTAAAGAACTATATGTGCCAGAAATAAATTCATTTGGCAAAAAACAGATGCGGAAC 813
QY 729 TCTCCCTTCAATTAATTAAGCTGATCTCCAGTGGTCCAGGTTTGTAGCGAAGG 788
Db 814 ATTACAGCACTGAAATTAATTAAGTGGATATGCACTGGGTGAGGTTTGGCAGAGG 873
QY 789 CTGGGCACTCCCTCAAGGTTTTCATGCGGGGAGAGAGTACTTACAGGTTTATGCACTT 848
Db 874 TTGGCAACCCCATTTGAATGGCTTTATGACAGAGAGAGGAGTACTTTCAGTGGCTTTCATT 933

Db 94 GAGCAATAACGGCAGAACTGGGAATCGAGAGACCAACCAATGTCACCTACCAAGCCCA 153
Qy 69 CCATGTGACGAGGAATAAGAGAGGCAAGTGGTGGAAACAAGGGGTGGGTTCCGAGGATG 128
Db 154 TCATGTCAGCAGGAAACAAGAGAGTCAAGTGGTGGGACCAAGAGTGGCTTCGTGGTTG 213
Qy 129 TACCGTGTGCTAACAGAGTCTCTGGGTGCTGGAAACAACAGGATAAGTTTGGCCCTGGA 188
Db 214 CACAGTTTGGCTAACAGGCTTGTCTGGAGCGGGAAGACTACTGTGAGCATGGGCTTGG 273
Qy 189 GGAGTACCTTGTCTCCCATCCATCCCTTGTACTCCCTGGATGGGACAATGTCCGTC 248
Db 274 GGAGTACCTGCTTGTCTCATGTAATCCATGCTACACTCTGATGGTGAATATTCGTCA 333
Qy 249 TGGCCTTAAACAGAAATCTCGGATTCCTCTGGGACAGAGAGGAAATATCCCGCGAT 308
Db 334 AGGTCTCAATAAATAATCTGGCTTTAGTCTCTGAAACAGAGAGAGAAATGTTCGACGCAT 393
Qy 309 TGCTGAGTGGCTAAGCTGTTTGTCTGATGCTGGTCTGGTCTGCATTAACAGCTTTATTC 368
Db 394 CGCAGAAGTTGCTAAACTGTTTGCAGATGCTGGCTTAGTGTGCATCAAGTTTCATATC 453
Qy 369 TCCATTCCGAAAGGATCGTGAAGATGCCCGCAAAATACATGAATCAGCAGGCTGCCATT 428
Db 454 ACCTTACACTCAGGATCGCAACATGCAAGCAATTCATGAAGGTGCAAGTTTACCGTT 513
Qy 429 CTTTGAATATTGTAGATGACCTCTAAATATTGTGTAAGCAGAGACAGTAAAGGCC 488
Db 514 TTTTGAAGTATTGTGATGCTCTCTGCAATGTTGTGAACAGAGGATGTCAAGGACT 573
Qy 489 CTATAAAGGGCCAGAGCTGGGAGATTAAAGGATTTACAGTATTGATCTGATTATGA 548
Db 574 CTACAAAAGCCGCGCAGGAGAAATTAAGGTTTCACTGGGATCGAATCTGAATATGA 633
Qy 549 GAAACCTGAAACTCCTGAGGCTGTCTTAAACCAATTTGTCCACAGTGAAGTCTGT 608
Db 634 AAAGCCAGAGGCCCTGAGTGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 693
Qy 609 CCACCAAGTGTGAACTTCTGCAAGAGCAGAACTTGTACCTTACTATATCAAGA 668
Db 694 CCACCAAGTGTGAACTTCTGCAAGAGCAGAACTTGTACCTTACTATATCAAGA 753
Qy 669 TATCCAGCAACTCTTGTGCGGAAACAACTTGTACCACTGCTGCTGCTGCTGCTGCT 728
Db 754 AGTAAAGAACTATATGTCAGAGAAATAAATCTTATTTGGCAAAACAGATGCGGAAC 813
Qy 729 TCTCCCTTCATATCAATTAAGTGTGATCTCCAGTGGGTGAGGCTTTGAGCGAAG 788
Db 814 ATTACCAAGCACTGAAATTAATTAAGTGGATATGCACTGGGTGAGGCTTTTGGCAGAAG 873
Qy 789 CTGGCCCACTCCCTCAAGGTTTCATGCGGGAGAGGAGTACTTACAGGTTATGCACTT 848
Db 874 TTGGCAACCCCATTAAGTGGCTTTATGAGAGAGGAGGAGTACTTGCAGTGCCTTCATTT 933
Qy 849 TGACACCCCTGCTAGATGATGCGGTGATCAACATGAGCATCCCATTTCTACTGCCGPTC 908
Db 934 TGATTGCTTCTGATGAGGTGCTATTAATCTGCTAGTACCTATATCTTCTGACTGCGAC 993
Qy 909 TGCAGAGGATAAGACAGCGGTGAGGAGGTGAGCAAGTTTGTCTGTCACATGGTGGAGC 968
Db 994 TCATGAAGATAAAGAGAGGCTGAGCGCTGTACAGCATTTTGTCTGATGATGAGGSCCG 1053
Qy 969 GAGGTAGCTATCTTACGACAGCTGAATCTATGAACACAGAAAGAGAAAGCGGTTC 1028
Db 1054 CCGTGTGGCCATCTTCTGCAATCCAGAGTTTGTGAGCAGAGAAAGAGGAGCGGTGTC 1113
Qy 1029 CCGTGTGTGGGACAACTGTCAAAACACCCCATATCAAAATGCTGATGAAAGTGG 1088
Db 1114 CAGACAGTGGGACAGACATGCAAGAACACCCCTATATTAAGTGGTGAAGCAAG 1173
Qy 1089 GGACTGGCTGGTGGGACACCTTCAAGTGTGCTGGAGAAATTAAGATGAATGAGGCT 1148
Db 1174 AGATTGGCTGATTGGGAGGATCTTCAAGTCTTGGATCGAGTTTATTTGGAAATGATGCT 1233

Qy 1149 GGACCAATACCCTCTGACACCTCTGAGCTCAACACAGAAATGTAAAGAAATGAATGCTGA 1208
Db 1234 TGATCAGTATCGTCTTACTCTCTACTCTAAAGCAGAAATTTAAAGATATGAATGCTGA 1293
Qy 1209 TGGCGTGTTCGATTCGATTCGCAATCCCTGTCCCAATGGCATGCCCTGTTCATGCA 1268
Db 1294 TGCTGTCTTTCGATTTCAACTACGCAACCCAGTGCACAATGGACATGCTTGAATGCA 1353
Qy 1269 GGACACCTGCCCGAGGCTCTAGAGAGGGGTACAGCACCCGGTCTCTTACTACTACACC 1328
Db 1354 GGATACCAATAGCAACTTCTAGAGAGGGGTACCGGGCCCTGTCTCTCTCTCCACCC 1413
Qy 1329 TCTGGCGGCTGGACCAAGGATGAGATGTCTCTAGACTGGGATGAAGACAGCAGC 1388
Db 1414 TCTGGTGGCTGGCAAAAGGATGACGATGTTCTTGTATGTGGGTATGAAGCAGCATG 1473
Qy 1389 GGCTGTCTCGAGGAAGGGTCTGGATCCCAAGTCAACCAATTTTGGCCATCTTCCGTC 1448
Db 1474 TGCAGTGTGGAGGAAGGTTCTGAATCTCTGAGACGACAGTGGTGGCCATCTTCCCATC 1533
Qy 1449 TCCCATGTATATCTGTGCCCCACAGAGGTCCAGTGGCACTGCAAGTCCCGGATGATTC 1508
Db 1534 TCCCATGATGATGCTGACCAACTGAGTCCAGTGGCATTCGAGACGCGATGTTGC 1593
Qy 1509 GGTGCGCAATTTTACATTTGGGGAGGAGCCCTGCAGGAATGCCCATCTCTGAACCAA 1568
Db 1594 AGGAGCAACTTTTACATTTGGAGCAGAGACCTGTGGCATGCTCATCCAGAAACAGG 1653
Qy 1569 GAAGGATCTGTATGAACCCACTCATGGGGCAAGTCTTGAAGCATGGCCCTGCGCTCAC 1628
Db 1654 GAAGATCTTTATGAGCCAACTCATGTTGGCCAAAGTCTGACGATGGCCCTGCTTAT 1713
Qy 1629 CTCTGTGGAATCATCTCCATTCGAGTGGCTGCCTACAAAGAGCCAAAAGCCATGGA 1688
Db 1714 CACTTTGGAATAGTTCCTTTCCAGTTGCACTTTACAAAGAAAGAGCGTATGGA 1773
Qy 1689 CTCTATATTCACGACGAGCAGATGATTTGACTTTCATCTCAGGAATCGAATGAGAA 1748
Db 1774 CTACTATGACTTGAACCACTGAAGACTTTGAATTTATTCAGGAACACGAATGCGCA 1833
Qy 1749 GCTGCCCGGAGAGAGATCCCGAGATGGCTTCATGCCGCCCAAGCATGGAAGGT 1808
Db 1834 ACTTGCTCGAGAGGCCAGAAACCTGAGGTTTCATGGCTCCCAAGGCTTGGACGT 1893
Qy 1809 CCGTACAGATTTATACAGTCCCTTGGAGAA 1838
Db 1894 GCTGACAGAATACTACAAATCCTTGGAGAA 1923

RESULT 5

US-09-966-262-116
; Sequence 116, Application US/09966262
; Publication No. US20030050461A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 87 Human Secreted Proteins
; FILE OF INVENTION: P2004P1
; CURRENT APPLICATION NUMBER: US/09/966,262
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: US 09/154,707
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: PCT/US98/05311
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: US 60/041,277
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/042,344
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,276
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/041,281
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: US 60/048,094

Db 694 CCAGCAAGTGTGGAACTTCTACAGGAACGGGATATGTACTGTGGATGCATCTTATGA 753
 QY 669 TATCCAGCAACTCTTTGTGCGGGAACAAACCTTCACCAAGCTCCAGCTGAGCGTGAAC 728
 Db 754 AGTAAAGCACTATATGTCCAGAAATAAATCTTATTTGSCAAAACAGATGCGGAAC 813
 QY 729 TCTCCCTTCATATCAATTAAGTCAAGTCTCCAGTGGGTCCAGGTTTTTGGAGCAAG 788
 Db 814 ATTACGAGCACTGAAATTAATAAAGTGGATGCGTGGGTGCGAGTTTGGCAGAAG 873
 QY 789 CTGGGCCACTCCCTCAAGGTTTCATCGGGAAGAGGAGTACTTACAGGTTATGCATTT 848
 Db 874 TTGGCAACCCATTAATGAGCTTTATGAGAGAGAGGAGTACTTGCAGTGCCTCATTT 933
 QY 849 TGACACCTGTGTAGATGCGGTGATCAACATGAGCATGCCCATTTGACTGCCGCTC 908
 Db 934 TGATTGTCTCTGGATGGAGGTGTCATTAACCTTGTCACTACCTATAGTTCTGACTGCGAC 993
 QY 909 TCCAGAGTAAAGACACCGCTGGGAAGGTGAGCAAGTTTTCCTGGGCACATGGTGGAG 968
 Db 994 TCATGAAGATAAAGAGAGCGCTGGACGGCTGTACAGCATTTGCTCTGTATGAGGCGG 1053
 QY 969 GAGGTAGCTATCTTACGAGCGCTGAATTTCTATGACACAGAAAGAGGAGCGTGTTC 1028
 Db 1054 CCGTGTGCCATCTTCGCAATCCAGAGTTTTTTCAGCAGAGGAAGAGGCGGTGTGC 1113
 QY 1029 CCGTGTTTGGGGACAACATGTACAAAACACCCCATATCAAAATGGTGTGAAGTGG 1088
 Db 1114 CAGACAGTGGGAACGACATCAAGAACCCCTATATTAAGATGGTGTGACACAGG 1173
 QY 1089 GAGTGGCTGTGTGGAGACCTTTCAGGTGCTGGAGAAATAAGATGAATGATGGGCT 1148
 Db 1174 AGATTGGCTGTATGGAGGAGATCTTCAAGTCTTGGATCGAGTTTATTTGGAATGATGGTCT 1233
 QY 1149 GACCAATACCTGTGACACCTCTGGAGCTCAACAGAAATCTAAGAAATCAATGCTGA 1208
 Db 1234 TGATCAGTATCTGCTTACTTCTACTGAGCTAAGCAGAAATTAAGATATGAATGCTGA 1293
 QY 1209 TCGGTGTGTGATTCAGTTGCGCAATCTCTCCACAAATGCCATGCCCTGTTGATGCA 1268
 Db 1294 TGCTGTCTTGTCAATTCACATACGCAACCCAGTGCACATGACATGCCCTGTTAATGCA 1353
 QY 1269 GACACCTGCGCAGCGCTCTAGAGAGGGCTACAGCACCCGGTCTCTACTACACCC 1328
 Db 1354 GATACCCATTAAGCAACTCTAGAGAGGGGTACCGGGCGCTGCTCTCTCCCTCACCC 1413
 QY 1329 TCTGGCGCTGGACCAAGATGACGATGTGCTCTAGACTGGCGGATGAAGCAGCAGCG 1388
 Db 1414 TCTGGTGTGGACAAAGGATGACGATGTCTCTTGTGTCGCTATGAAGCAGCATGC 1473
 QY 1389 GGCTGTGCTGAGGAAGGGTCTCGATGCCAAGTCAACCATTTGTCGCTATCTTTCCGTC 1448
 Db 1474 TGCAGTGTGGAGGAAGGATGCTGAATCTCTGAGACGACAGTGGTGGCCATCTTCCCATC 1533
 QY 1449 TCCCATGTTATGCTGCGCCACAGAGTCCAGTGGCAGTGCAGTCCCGATGATGTC 1508
 Db 1534 TCCCATGATGATGCTGCGACCAACTGAGTCCAGTGGCATTCGACAGCAGGATGGTTC 1593
 QY 1509 GGGTGCCAAATTTCTACATTTGGGAGGAGCCCTGCAGGAATGCCCATCTCTGAACCAA 1568
 Db 1594 AGGACCAACTTTTACATTTTGGACGAGACCCCTGCTGGCATGCTCATCAGAAACAGG 1653
 QY 1569 GAAGATCTGATGAACCACTCAATGGGGCAAGGTCTTGAGCATGGCCCTGGCCCTCAC 1628
 Db 1654 GAAGATCTTTATGAGCAAGTCAATGTTGCCAAGTGTGACGATGTCGAGTGGTGGTTAAT 1713
 QY 1629 CTCTGTGAATCATTTCCATTCGAGTGGCTGCTGCTACACAAAGCCAAAGCCATGGA 1688
 Db 1714 CACTTGTGAATGATGCTTTCGAGTTTGCAGCTTACAAAGAAAGAAAGACGCTATGGA 1773
 QY 1689 CTCTGATGATCCAGCAAGCAATGAGTTTGTGACTTCTCAGGAAGTCAATGAGGAA 1748
 Db 1774 CTACTGACTCTGACACCATGAGACTTTGAATTTATTTTCAGGAACCAAGATGCGCAA 1833

QY 1749 GCTGCGCCGGAAGAGAGAAATCCCCAGATGCTTCTATGCGCCCCCAAGCATGGAAGT 1808
 Db 1834 ACTTCTCGAGAAGCCAGAAACACCTGTAAGGTTTCATGCTGCCAGGCTCCCAAGGCTTGGACCGT 1893
 QY 1809 COTGACAGATTAATACAGTCCCTCCCTGGAGAA 1838
 Db 1894 GCTGACAGAATACTACAAATCCTTGGAGAA 1923

RESULT 7
 US-09-983-966-116
 ; Sequence 116, Application US/09983966
 ; Publication No. US20030060619A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young et al.
 ; TITLE OF INVENTION: 87 Human Secreted Proteins
 ; FILE REFERENCE: P2004PI
 ; CURRENT APPLICATION NUMBER: US/09/983,966
 ; CURRENT FILING DATE: 2001-10-29
 ; PRIOR APPLICATION NUMBER: 09/154,707
 ; PRIOR FILING DATE: 1998-09-17
 ; PRIOR APPLICATION NUMBER: PCT/US98/05311
 ; PRIOR FILING DATE: 1998-03-19
 ; PRIOR APPLICATION NUMBER: US 60/041,277
 ; PRIOR FILING DATE: 1997-03-21
 ; PRIOR APPLICATION NUMBER: US 60/042,344
 ; PRIOR FILING DATE: 1997-03-21
 ; PRIOR APPLICATION NUMBER: US 60/041,276
 ; PRIOR FILING DATE: 1997-03-21
 ; PRIOR APPLICATION NUMBER: US 60/041,281
 ; PRIOR FILING DATE: 1997-03-21
 ; PRIOR APPLICATION NUMBER: US 60/048,094
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,350
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,188
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,135
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/050,937
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,187
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,099
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,352
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,186
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,131
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,096
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,355
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,160
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,351
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/048,154
 ; PRIOR FILING DATE: 1997-05-30
 ; PRIOR APPLICATION NUMBER: US 60/054,804
 ; PRIOR FILING DATE: 1997-08-05
 ; PRIOR APPLICATION NUMBER: US 60/056,370
 ; PRIOR FILING DATE: 1997-08-19
 ; PRIOR APPLICATION NUMBER: US 60/060,862
 ; PRIOR FILING DATE: 1997-10-02
 ; NUMBER OF SEQ ID NOS: 343

Matches	1305;	Conservative	12;36;	Pro: NO. 0;	Mismatches	525;	Indels	0;	Gaps	0;
QY	9	GATCAAGAAGCAAAAGACGAGAACCCACAGCAAGAAATCCACCAATAGTCTACTACAGGCCCA	68							
Db	94	GAGCAATAACGCGCAGAACTGGGGAATGCGAGAGAACCAATGCTCACTTACCAAGGCCCA	153							
QY	69	CCATGTGACGAGCAANTAAAGAGAGGGCAAGTGGTTGGAAACAAGGCGGTTCGAGGATG	128							
Db	154	TCATGTGACGAGCAACAAGAGAGTCAAGTGGTGGGACGAGAGGTGGCTTCGTGGTGTG	213							
QY	129	TACCGTGTGCTTAACAGGTCTCTCTGGTCTGGAAACAACAGATAAGTTTGCCTTGGGA	188							
Db	214	CACAGTTTGGCTTAACAGGCTTGTCTGGAGCGGGAAGACTACTGTGAGCATTGGCCTTTGGA	273							
QY	189	GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCTCGTGATGGGACAAATGTCCTGCA	248							
Db	274	GGAGTACCTGGTTTGTCAATGATTCATCTACACTCTGGATGGTGACAAATATTCGTCA	333							
QY	249	TGGCCTTAAACAGAAATCTCGGATCTCTCTGGGACAGAGAGAAATATCCGCCGAT	308							
Db	334	AGGTCTCAATATAAAATCTTGGCTTTAGTCTCTGAAGACAGAGAGAATGTTTCGACGAT	393							
QY	309	TGCTGAGGTGGCTTAAGCTGTTTCTGTAGTCTGGTCTGCTGCAATACCAGCTTTATTTC	368							
Db	394	CGCAGAAGTTGCTTAAACTGTTTCGATGCTGGCTTGTGTGTCATCAAGTTTCATATC	453							
QY	369	TCCATTTCGAAAAGATCGTGAGAAATGCCGCAAAATACATGAATCAGCAGGGCTGCCATT	428							
Db	454	ACCTTACACTCAGGATCGCAACAATGCAAGGCAAAATTCATGAAGTGCAGTTTACCCTT	513							
QY	429	CTTTGAAATATTGTAGTCACTCTTAAATATTGTGAAAGCAGAGACGCTTAAAGGCT	488							
Db	514	TTTTTGAAGTATTGTTGATGCTCTCTGCAATGTTTGTGTAACAGAGGATGTCAAAGGACT	573							
QY	489	CTATAAAGGGCCACAGACTGGGAGATTAAGAGATTTACAGGTATTGATCTGTATATGA	548							
Db	574	CTACAAAAAGCCGGGACGAGAAATTAAGGTTTCACTGGGATCGATTCTGTAATATGA	633							
QY	549	GAAACCTGAAACTCCTCAGCGTGTGCTTAAACCAATTTGTCCACAGTGAGTGACTGTG	608							
Db	634	AAAGCCAGAGGCCCTTGAGTTGGTGTCTGAAACACAGACTCCTGTGATGTAAATGACTGTG	693							
QY	609	CCACCAGGTAGTGGAACTTCTGCAAGACGACAGAACTTGTACCCCTTACTATAATCAAGA	668							
Db	694	CCAGCAAGTTTGGAACTTCTACAGGAACGGATATTGTACCTGTGGATGCATCTTATGA	753							
QY	669	TATCCAGAACTCTTTGTGGCGGAAACAACATTGACACGCTCCGAGCTGAGGCTGAAC	728							
Db	754	AGTAAAAGAACTATATGTGCGCAAAAATAAATCTTATTTGCAAAAAACAGATGCGGAAC	813							
QY	729	TCCTCCCTTCATTATCAATTACTAAGCTGGATCTCCAGTGGGTCCAGGTTTGTAGCGAAG	788							
Db	814	ATTACACGCACTGAAAAATTAATAAGTGGATATGCACTGGGTGCGAGTTTGGCAGAGG	873							
QY	789	CTGGGCACTCCCTCAAAGGTTTCAATGCGGGAGAGGAGTACTTACAGGTTATGCACTT	848							
Db	874	TTGGGCNACCCATTGNAATGGCTTTATGACAGAGAGGAGTACTTTCAGTGCCTTCATT	933							
QY	849	TGACACCTGCTAGATGATGGCTGTATCAACATGAGCATCCCCATTGTACTGCCCGTCTC	908							
Db	934	TGATTTGCTCTCGATGGAGGTGTCATTAACTTGTCTAGTACCTATAGTCTGACTGCGAC	993							

GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPL300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIORITY FILING DATE: 2001-08-24
PRIORITY FILING DATE: 2000-08-24
PRIORITY FILING DATE: 2000-08-24
PRIORITY FILING DATE: 2001-01-16
PRIORITY FILING DATE: 2001-01-16
PRIORITY FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2084
LENGTH: 1398
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2084

Query Match 19.5%; Score 359.8; DB 9; Length 1398;
Best Local Similarity 58.7%; Pred. No. 3e-111;
Matches 682; Conservative 0; Mismatches 467; Indels 12; Gaps 3;

676 GAATCTTTGTCGCGGAAACAACTTGACACGTCGCGAGCTGAGGCTGAAACTCTCCCT 735
184 GATCTTTGTTACCGGAACCGAGAGCGGAGAGAAACACAGAGCGCGGATTTGGCG 243
736 TCATTATCAATTAAGTGGATCTCCAGTGGTCCAGGTTTTCAGCGAAGGCTGGGCC 795
244 AGAGTGAGATTGACGGCGATTGATTGCAATGGATGATGTTGAGTGAAGGTTGGGCT 303
796 ACTCCCTCAAGGTTTCATCGCGGAGAGAGTACTTACAGGTTATGCACTTTGA ---- 851
304 AGTCCCTCTGCTGGGTTATGAGGGAATCTGAGTTCCTCCAAACTCTTCATTTCAATTG 363
852 --CACCTGCTGATGATGCGGTGATCAACATGAGCATCCCATTTGACTCCCGTCTCT 909
364 TTGAATCTGATGGTCTGTTGTTAATATGCTGTGCGCTATGTTCTTGCATGAT 423
910 GCAGAGGATAAGACACGCGCTGGAAGGTGCGAGCAAGTTTGCCTGGCACATGTTGGAGG 969
424 GATCAACAAAAGCCCTAATCGGTGAATCTAAACGTGTCTCCCTGTTGATTTCTGATGAT 483
970 AG---GGTAGTATCTACGAGACGCTCAATTCTATGAACACAGAAAGAGACGCTGT 1026
484 AATCAATCGCTATTCTCAATGATATTGAGATTTTATAAATCCGAAAGAGAGCGAATA 543
1027 TCCGCTGTTTGGGGGACAAATGTACAAAACACCCCATATCAAAATGGTGAT---GGAA 1083
544 CGAGAACTTGGGTACGACTGCACCGGGTTGCTTATGTAGAGAGCGCATACCAAT 603
1084 AGTGGGACTGGCTGTTGTTGGAGACTTCCAGTGTGGAGAAATAAGATGAATGAT 1143
604 GCTGGAGACTGGCTCAATGGGGGTGATCTTGGGTTTGGAAACCTGTTAAGTACAATGAT 663
1144 GGGCTGGACCAATACCGTCTACACCTCTGAGCTCAACAGAAATGTAAGAATGAT 1203
664 GGGCTGATCGTTTTCAGGCTTTCCTCGGTTTGAACCTGCGTAAAGGAGCTAGAGAAACGTTG 723
1204 GCTGATCGGTGTTTGGATTCAGTTCGCAATCTGTCCCAATCTGTCCCAATGCGCTGTTG 1263
724 GCGATGCGGTCTTGGCTTTCAGCTTAGGAACCCAGTTTCAATATGACATGCTCTCTT 783
1264 ATGAGGACACTGCGGAGGCTCTTAGAGAGGGGCTACAGACCCCGCTCTCTCTACTA 1323
784 ATGACTGATCTCTGAGAGACTTCTTGGATGGTTTATAAAACCCCTATCTCTTTGCTT 843
1324 CACCTCTGGCGGCTGGACCAAGGATGACGATGCTCTAGACTGGCGGATGAAGCAG 1383

Db 844 CATCATTTGGAGGGTTTACAAAAGCTGATGATGTTCTCTTAAGCTGCGGAATGAACAG 903
Qy 1384 CACGCGGCTGCTCGAGGAAGGGTCTCTGATCCCAAGTCAACCATTTGTTGCCATCTTT 1443
Db 904 CACGAGAAGTCTAGAGGATGTTCTTCTGATCCAGAGACTACTGTGTTTCCATATTC 963
Qy 1444 CCGTCTCCCATGTTATATGCTGGCCCCACAGAGGTCCAGTGGCACTGCAAGTCCCGGATG 1503
Db 964 CCATCTCCAATGCTCTATGCTGGTCCAAACCGAAGTCCCAATGCGACGCAAGGCTAGGATC 1023
Qy 1504 ATTGGGGTGCATTTCTACATTTCTGGGGAGGACCTTCGAGGATGCCCATCTCTGAA 1563
Db 1024 AATGCTGGTAACTTCTACATTTGCTGGTAGGATCGGGTGGATGGGTCAATCCCGTG 1083
Qy 1564 ACCAAGAAGGATCTGATGAACCCACTCATGGGGCAAGGTCTTTCAGCATGGCCCTGGC 1623
Db 1084 GAGAAACGTGATCTGACGATGCTCATCAGGGAAGAAGTCTTAAGCATGGCTCTCGA 1143
Qy 1624 CTCACCTCTGTGGAAATCATTCATTCGAGTGGCTGCTACACAAAGCCAAAGGCC 1683
Db 1144 CTCGAACGACTCAACATTTCTTCCCTTTCAGGGTTCAGGCTTCTTTCATTTCTGGCAGTAA 1203
Qy 1684 ATGCACTTCTATGATCCAGCAAGGCACAAATGAGTTTGACTTTCATCTCAGGACTCGAATG 1743
Db 1204 ATGGCTTCTTTGATCCCTCAAGGGCTCAGGACTTCTTTCATTTCTGGCAGTAAAGT 1263
Qy 1744 AGAAGCTCGCGGAGAGAGAAATCCCGCAGATGCTTCATGGCCCCCAAGCATGG 1803
Db 1264 AGACATTTGGCAAGACAGAGAAACCCACAGATGATTTATGTGCTCGGAGGCTGG 1323
Qy 1804 AAGTCTCTGACAGATTATTAC 1824
Db 1324 AAGTCTCTGTGATTACTAC 1344

RESULT 11

US-09-880-107-360/c
Sequence 360, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 360
LENGTH: 574
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA165526
NAME/KEY: unsure
LOCATION: (1)..(574)
OTHER INFORMATION: n = a or c or g or t
US-09-880-107-360

Query Match 17.1%; Score 315; DB 10; Length 574;
Best Local Similarity 73.9%; Pred. No. 2.9e-96;
Matches 424; Conservative 0; Mismatches 148; Indels 2; Gaps 2;

Qy 1155 ATACCGTCTGACACTCTGGAGCTCAACAGAAATGTAAGAAATGAATGCTGATCGGT 1214
Db 574 ANATCGTTTACTCTTACTGAGCTAAACCAAGAAAT-TNAGGTTTGAATGCTGATGCTGT 516

1215	QY	GTTTGGCAATCCAGTTGGCGCAATCCTGTCCACAATGGCCATGCCCTGTTGATGCAGGACAC	1274
515	Db	CTTTGGCAATTTCAACTAGCCACCCCGACATGACATGGACATGCCCTGTTAATGCAGGATAC	456
1275	QY	CTGCCGAGGCTCCTAGAGAGGGGCTACAGCACCCGGTCTCTCTACTACACCTCTGGG	1334
455	Db	CCATAAGCAACTTCTAGAGAGGGGCTACCGCGCCCTGTCTCTCTNCITCCCTCTGGG	396
1335	QY	CGGCTGGACCAAGATGACCATGTGCCCTCTAGACTGSCGGATG-AAGCACACGCGGCTG	1393
395	Db	TGGCTGGACAAAGATGACCATGTCTTTGATGTGGCGTATGAAGACGACATGCTGCAG	336
1394	QY	TGCTCGAGGAAGGGTCTTGATCCCAAGTCAACCATTTGTGGCATCTTTCCCTCTCCCA	1453
335	Db	TGTTGGAGGAAGAGTTCTGAATCCTGAGACGACAGTGGTGGCATCTTCCCATCTCCCA	276
1454	QY	TGTTTATGCTGCCCCACAGAGGTCAGTGGCACTGCGAGTCCGGATGATTCGCGGTG	1513
275	Db	TGATGTATGCTGGACCAACTGAGTGCCAGTGGCATTCGACAGCACGGATGGTTGCAGGAG	216
1514	QY	CCAATTTTACATTTGTGGGAGGAGCCCTGACGGAATGCCCATCTCTGAACCAAGAGG	1573
215	Db	CCAACTTTACATTTGTGGACGAGACCCCTGCTGGCATGCCATCCAGAAACAGGGAAGG	156
1574	QY	ATCTGTATGAACCCACTCATGGGGGGAAGGTCCTTGAGCATGGCCCTGGCCTCACCTCTG	1633
155	Db	ATCTTTATGAGCAAGTCATGGTGCCAAAGTCTGACGATGGCCCTGGTTTAACTACATT	96
1634	QY	TGGAATCATTTCCAGTGGGTCCTTACAAACAAAGCCAAAAAGCCATGGACTTCT	1693
95	Db	TGGAATAGTTCCCTTTCGAGTTGCAGCTTACAAACAAGAAAAAGAGCGTATGGACTACT	36
1694	QY	ATGATCCAGCAAGGCACATGAGTTTGACTTCAT	1727
35	Db	ATGACTCTGAACACCATGAAGACTTTTGAATTTAT	2

RESULT 12

US-09-796--692-8898
Sequence 8898, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
OF
FILE REFERENCE: 2077.001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378

```

; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8898
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (572)
; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-8898

Query Match      13.2%; Score 243; DB 9; Length 600;
Best Local Similarity 75.2%; Pred. No. 1.1e-71;
Matches 303; Conservative 0; Mismatches 100; Indels

Qy      1436  CCATCTTTTCGCTCCCCATGTTATATGCTGCGCCCCACAGAGAGTCCAGCTGG
Db      1     CCATCTTTCCCATCTCCCATGATGATGCTGGACCAACTGAGGTCCAGTGG
.Oy      1496  CCGCGGATGTTGCGGGTCCCAATTTCTACATTTGTGGGAGGAGACCCCTGCAC
Db      61     CACGGATGGTTGCAGGAGCCCACTTTTACATTTGTGGACGAGACCCCTGCT
Qy      1556  ATCCTGAAACCAGAAGGATCTGTATGACCCCACTCATCTGGGGGCAAGGTC
Db      121    ATCCAGAAACAGGGAAGGATCTTTATGAGCCCAAGTCATGGTGCCCAAGTGA
Qy      1616  CCCTCGCCCTCACCTCTGTGAAATCATTTCCATTCGAGTGGCTGCCCTTAC
Db      181    CCCCTGGTTTANTCACTTTGGAAATAGTTCCCTTTTCGAGTTGCAGCTTAC
Qy      1676  AAAAGCCATGGACTTCTTATGATCCAGCAAGGCACAAATGAGTTTGACTTTC
Db      241     AGAAGCGTATGGACTACTATGACTCTGAACACCATGAAGACTTTTGAATTT
Qy      1736  CTGCAATGAGGAAGCTCGCCGGGAAGGAGAGAAATCCCCACAGATGGCTTTC
Db      301     CAGAAATGCGCAAACTTGTCTCGAAGAGGCCAAGAACCAACCTTGAAGGTTTC
Qy      1796  AAGCATGGAAGGTCCTCAGAGATTTATTACAGGTCCTCGTGAGAA 1838
Db      361    AGGCTTGGACCGTCTGCACAGAATACTACAATCTCTTGAGAA 403

```

RESULT 13

US-10-040-862-8898

Sequence 8898, Application US/10040862

Publication No. US20030078396A1

GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander

APPLICANT: Algate, Paul A.

APPLICANT: Mannion, Jane

APPLICANT: Retter, Marc

APPLICANT: Corixa Corporation

TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther

TITLE OF INVENTION: Hematological Malignancies

FILE REFERENCE: 014058-013520US

CURRENT APPLICATION NUMBER: US/10/040,862

CURRENT FILING DATE: 2001-11-06

PRIOR APPLICATION NUMBER: US 60/186,126

PRIOR FILING DATE: 2000-03-01

PRIOR APPLICATION NUMBER: US 60/190,479

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: US 60/200,545

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: US 60/200,303

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,779

PRIOR FILING DATE: 2000-04-28

PRIOR APPLICATION NUMBER: US 60/200,999

PRIOR FILING DATE: 2000-05-01

```
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8898
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (572)
; OTHER INFORMATION: n=A,T,C or G
US-10-040-862-8898
```

Query Match 13.2%; Score 243; DB 9; Length 600;

Best Local Similarity 75.2%; Pred. No. 1.1e-71;

Matches 303; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```
QY 1436 CCATCTTCCGTCCTCCATGTTATATGCTGGCCCCACAGAGTCCAGTGGCAGTGCAGT 1495
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 1 CCATCTTCCATCTCCCATGATGATGCTGGACCAACTGAGTCCAGTGGCAGTGCAGAG 60

QY 1496 CCGGATGTTGGGTGCGCAATTCTACATTTGGGGAGGACCTGTCAGGAATGCCCC 1555
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 CAGGATGTTGGAGAGGCACTTTTACATTTGGGACGAGACCTGCTGGCATGCCCTC 120

QY 1556 ATCTCGAAACCAAGAGGATCTGTATGAACCCACTCATGGGGGCAAGTCTTCAGCATGG 1615
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 ATCCAGAAACAGGAAGGATCTTTATGAGCCAAAGTCATGTCGCAAGTGTCTGACATGG 180

QY 1616 CCGCTGGCCCTACCTCTGTGGAAATCATTCATTCGGAGTGGCTGCCTACACAAAGCCA 1675
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 CCGCTGGTTTATCACTTTGGAATAGTTCCCTTTTCGAGTTGAGCTTACACAAAGAAA 240

QY 1676 AAAAGCCATGGACTTCTATGATCCAGCAAGGCACAAATCAGTTTGACTTCATCTCAGAA 1735
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 AGAAGCGTATGACTACTATGACTCTGAACACCATGAAGACTTTGAATTTATTCAGGAA 300

QY 1736 CTCGAATGAGGAAGCTCGCCGGGAAGGAGAGAAATCCCCAGATGGCTTCATGGCCCCCA 1795
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 CAGCAATGCGCAAACTGTCTCGAAGAGGCCAGAAACCACTGAAGGTTTCATGGCTCCCA 360

QY 1796 AGCATGGAAGTCTCTGACAGATTTATACAGTCCCTCGAGAA 1838
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 AGGCTTGGACCGTCTGACAGAAATACATAAATCCTTTGGAGAA 403
```

RESULT 14

US-10-198-846-5659

; Sequence 5659, Application US/10198846

; Publication No. US20030099974A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF BREAST CANCER

; FILE REFERENCE: MRI-049

; CURRENT APPLICATION NUMBER: US/10/198,846

```
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 803
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 620, 658, 687, 703, 727, 754, 764, 765, 783, 792
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5659
```

Query Match 12.7%; Score 233.4; DB 9; Length 803;

Best Local Similarity 99.6%; Pred. No. 2.4e-68;

Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1475 AGTCCAGTGGCACTGCAGGTCCCGGATGATTCGGGGTGCCAAATTTCTACATTTGGGGA 1534
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 AGTCCAGTGGCACTGCAGGTCCCGGATGATTCGGGGTGCCAAATTTCTACATTTGGGGA 186

QY 1535 GGGACCTCGAGGAATGCCCATCCCTCTGAAACCAAGAGGATCTGTATGAACCCACATCATG 1594
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 GGGACCTCGAGGAATGCCCATCCCTCTGAAACCAAGAGGATCTGTATGAACCCACATCATG 246

QY 1595 GGGGCAAGTCTTTGACCATGGCCCTGGCTCACCTCTGTGGAATCATTTCCATTCGGAG 1654
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 GGGGCAAGTCTTTGACCATGGCCCTGGCTCACCTCTGTGGAATCATTTCCATTCGGAG 306

QY 1655 TGCTGCCTACACAAAGCCAAAGCCATGACTTCTATGATCCAGCAAGCA 1709
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 TGCTGCCTACACAAAGCCAAAGCCATGACTTCTATGATCCAGCAAGCA 361
```

RESULT 15

US-10-198-846-12859

; Sequence 12859, Application US/10198846

; Publication No. US20030099974A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF BREAST CANCER

; FILE REFERENCE: MRI-049

; CURRENT APPLICATION NUMBER: US/10/198,846

; CURRENT FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/306,220

; PRIOR FILING DATE: 2001-07-18

; NUMBER OF SEQ ID NOS: 14084

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12859

; LENGTH: 826

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17,

; LOCATION: 18, 19, 824, 825, 826

; OTHER INFORMATION: n = A,T,C or G

US-10-198-846-12859

Query Match

Best Local Similarity 12.7%; Score 233.4; DB 9; Length 826;

Matches 234; Conservative 99.6%; Pred. No. 2.5e-68;

Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1475 AGTCCAGTGGCACTGCAGGTCCCGGATGATTCGGGGTGCCAAATTTCTACATTTGGGGA 1534
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 371 AGTCCAGTGGCACTGCAGGTCCCGGATGATTCGGGGTGCCAAATTTCTACATTTGGGGA 430
```

Qy 1535 GGGACCCCTGCAGGAATGCCCCATCCTGAAACCAAGAGGATCTGTATGAACCCACTCATG 1594
 Db 431 GGGACCCCTGCAGGAATGCCCCATCCTGAAACCAAGAGGATCTGTATGAACCCACTCATG 490
 Qy 1595 GGGGCAAGGTCTTGAGCATGGCCCTGGCCCTCACCCTCTGTGGAAATCATTCCATTCCGAG 1654
 Db 491 GGGGCAAGGTCTTGAGCATGGCCCTGGCCCTCACCCTCTGTGGAAATCATTCCATTCCGAG 550
 Qy 1655 TGGCTGCCTACAAACCAAGCCCAAAAGCCATGGACTTCTATGATCCAGCAAGGCA 1709
 Db 551 TGGCTGCCTACAAACCAAGCCCAAAAGCCATGGACTTCTATGATCCAGCAAGGTA 605

Search completed: June 4, 2003, 19:09:39
 Job time : 298 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2003, 15:10:31 ; Search time 2709 Seconds
(without alignments)
11030.149 Million cell updates/sec

Title: US-09-898-165B-9

Perfect score: 1845

Sequence: 1 atgcggggatcaagaagca.....ggctccctggagaagaactaa 1845

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estnu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_est1.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mam.*

24: em_gss_mus.*

25: em_gss_others.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	907	49.2	956	14	BQ957456
2	905.6	49.1	1766	11	BC032513
3	866.2	46.9	951	9	AL540583
4	854.6	46.3	886	9	AL550460
5	853.6	46.3	915	14	BQ924373
6	777.6	42.1	872	14	BQ220909

7	757.2	41.0	860	13	BI831135
8	743.6	40.3	949	14	BQ926758
9	734.8	39.8	1041	13	BI829344
10	722.6	39.2	1054	14	BQ942867
11	704.8	38.2	874	14	BQ943498
12	672.4	36.4	917	14	BQ955947
13	629	34.1	683	12	BG484421
14	624.8	33.9	650	12	BE786139
15	624	33.8	727	13	BI103516
16	613.8	33.3	620	13	BM263737
17	582.8	31.6	1048	14	BM903600
18	580.4	31.5	873	14	BQ221395
19	579.4	31.4	684	12	BG438005
c	558.8	30.3	767	9	AI648997
20	546.8	29.6	972	14	BQ896549
c	528	28.6	640	12	BG823104
22	517.8	28.1	959	9	AL574314
23	513.2	27.8	906	12	BF182876
24	512	27.5	624	10	BE292722
25	508	27.5	944	13	BM458625
26	503	27.3	946	12	BG490797
27	488.8	26.5	492	14	BM708017
28	476.8	25.8	490	14	BQ776777
29	465	25.2	584	12	BQ723809
30	459	24.9	576	10	BE311856
31	441.4	23.9	870	13	BI101136
32	440.2	23.9	549	13	BI338434
c	436.2	23.6	557	10	AW571670
34	433.8	23.5	741	13	BI855664
35	425.8	23.1	709	14	BQ571984
36	424.2	23.0	886	14	BQ652788
37	424.6	23.0	938	12	BF788130
38	424.4	23.0	712	12	BG778039
39	423.6	23.0	1062	12	BG116418
40	419.2	22.7	979	13	BI257912
41	417.6	22.6	711	13	BI649417
42	415.6	22.5	804	13	BI852968
43	412.8	22.4	715	13	BI081933
44	407.2	22.1	916	9	AL517203
45					

ALIGNMENTS

RESULT 1
BQ957456
LOCUS BQ957456 956 bp mRNA linear EST 21-AUG-2002
DEFINITION AGNCOURT_8786574 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6371441
ACCESSION BQ957456
VERSION BQ957456
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 956)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DPGP/Gazdar
CDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium (LLNL) at: <http://image.llnl.gov>
Plate: LRCM2545 row: j column: 18
High quality sequence stop: 633.
Location/Qualifiers
1. .956

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6371441"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dr priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT      262 a      220 c      258 g      214 t      2 others
ORIGIN
Query Match
Best Local Similarity 98.18; Pred. No. 2.4e-255; Length 956;
Matches 938; Conservative 0; Mismatches 16; Indels 2; Gaps 2;

Qy 480 AAAAGGCTCTATAAAAGGCGCAGAGCTGGGAGATTAAAGGATTTACAGTATTGATTC 539
Db 1 AAAAGGCTCTATAAAAGGCGCAGAGCTGGGAGATTAAAGGATTTACAGTATTGATTC 60
Qy 540 TGATTATGAGAACTCAAACTCTGAGCGTGTCTTAAACCAATTTGCCACAGTGAG 599
Db 61 TGATTATGAGAACTCAAACTCTGAGCGTGTCTTAAACCAATTTGCCACAGTGAG 120
Qy 600 TGACTGTGTCACACGAGTAGTGAACCTTCTGCAAGAGCAGCAACATTTACCCCTATATAT 659
Db 121 TGACTGTGTCACACGAGTAGTGGAACTTCTGCAAGAGCAGCAACATTTACCCCTATATAT 180
Qy 660 AATCAAGATATCCACGAACCTCTTTGCGCGGAAACAAACTTGACACACGTCGCGAGCTGA 719
Db 181 AATCAAGATATCCACGAACCTCTTTGCGCGGAAACAAACTTGACACACGTCGCGAGCTGA 240
Qy 720 GGCTGAACCTCTCCCTTCATATCAATTAAGTCTGATGATCTCCAGTGGTCCAGGTTT 779
Db 241 GGCTGAACCTCTCCCTTCATATCAATTAAGTCTGATGATCTCCAGTGGTCCAGGTTT 300
Qy 780 GAGCGAAGGCTGGGCCACTCCCTCAAGAGTTTCATCGGGGAGAGGACTTACAGGT 839
Db 301 GAGCGAAGGCTGGGCCACTCCCTCAAGAGTTTCATCGGGGAGAGGACTTACAGGT 360
Qy 840 TATGCACTTTGACACCCCTGTAGATGATGCGGTGATCAACATGAGCATCCCCATTGTACT 899
Db 361 TATGCACTTTGACACCCCTGTAGATGATGCGGTGATCAACATGAGCATCCCCATTGTACT 420
Qy 900 GCCCGTCTCTGCAGAGGATAAGACACGCGCTGGAAGGTGCAGCAAGTTTCTCCTGGCACA 959
Db 421 GCCCGTCTCTGCAGAGGATAAGACACGCGCTGGAAGGTGCAGCAAGTTTCTCCTGGCACA 480
Qy 960 TGTGGACGAGGAGTACTTCTTACGAGACGCTGAATCTTATGAACACAGAAAAGGGA 1019
Db 481 TGTGGACGAGGAGTACTTCTTACGAGACGCTGAATCTTATGAACACAGAAAAGGGA 540
Qy 1020 ACCTGTTCCTGTTTGGGGACACATGTACAAACACCCCATATCAAAATGGTGAT 1079
Db 541 ACCTGTTCCTGTTTGGGGACACATGTACAAACACCCCATATCAAAATGGTGAT 600
Qy 1080 GGAAGTGGGACTGCTGCTGCTGAGACCTTCAGGTGCTGAGAAAATAGATGGAA 1139
Db 601 GGAAGTGGGACTGCTGCTGCTGAGACCTTCAGGTGCTGAGAAAATAGATGGAA 660
Qy 1140 TGATGGGCTGGACCAATACCGTCTGACCTCTGGAGCTCAACAGAAAATGTAAGAAAT 1199
Db 661 TGATGGGCTGGACCAATACCGTCTGACCTCTGGAGCTCAACAGAAAATGTAAGAAAT 720
Qy 1200 GAATGCTGATCGCGTGTTCATTCAGTTCGCAATCTGTCACAAATGGCCATGCCCT 1259
Db 721 GAATGCTGATCGCGTGTTCATTCAGTTCGCAATCTGTCACAAATGGCCATGCCCT 780

```

```

Qy 1260 GTTGATGCAGGACACCTGCCCGAGGCTCTAGAGAGGGGCTACAGCACCGGGTCTCTCT 1319
Db 781 GTTGATGCAGGACACCTGCCCGAGGCTCTAGAGAGGGGCTACAGCACCGGGTCTCTCT 840
Qy 1320 ACTACACCCCTCTGG-GGGCTGGACCAAGGATGAGATGTCCTCTAGACTGGCGGATGA 1378
Db 841 ACTACACCCCTCTGGGCGGCTGGACCAAGGATGAGATGTCCTCTAGACTGGCGGATGA 900
Qy 1379 AGCAGCACGGG-GCTGTGCTCGAGGAAGGGTCTCTGGATCCCAAGTCAACCATTTGT 1433
Db 901 AGCAGCACGGCTGGGCTTCCTCAAGAAAGGGTCTCTGGATCCCAAGTCAACCATTTGT 956

RESULT 2
BC032513 1766 bp mRNA linear HTC 27-JUN-2002
LOCUS Homo sapiens, Similar to 3'-phosphoadenosine 5'-phosphosulfate
DEFINITION Synthese 1, clone IMAGE:5492659, mRNA.
ACCESSION BC032513.
VERSION BC032513.1 GI:21619118
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1766)
JOURNAL Strausberg, R.
Direct Submission
Submitted (06-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@hri.nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgoe, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 69 Row: e Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 20127474
This clone has the following problem: frame shifted.

FEATURES
Location/Qualifiers
1..1766
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5492659"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_67"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"

BASE COUNT 486 a 366 c 476 g 438 t
ORIGIN
Query Match 49.1%; Score 905.6; DB 11; Length 1766;
Best Local Similarity 71.4%; Pred. No. 8.4e-255;
Matches 1193; Conservative 0; Mismatches 479; Indels 0; Gaps 0;

```

```
QY 9 GATCAAGAGCAAAAGACGGAGACACAGCAGAGAAATCCACCAATGTAGTCTATCAGGCCCA 68
   ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92 GAGCAATTAACGGCGAAGACTGGGGAATGCAGAGAGCAACCAATGTCACCTACCAAGCCCA 151
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 69 CCATGTGACGAGGAATAGACAGGCAAGTGGTGGGAACAAGGGGTGGGTTCCGAGGATG 128
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 TCATGTGACGAGGAACAAGAGGTCAGGTGGGGGACGAGGTGGCTTTCGTGGTTG 211
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 129 TACCGTGGGTAAACAGTCTCTCTGTGGTGGGAAACAAACAGTAAGCTTTTGCCCTGGA 188
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 212 CACAGTTTGGCTAACAGGCTTGTCTGGAGCGGGAAGACTACTGTGACATGGCCCTTGA 271
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 189 GGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGGAACAATGCCGTCA 248
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 272 GGAGTACCTGTTGTCTCATGTATTCATGCTACACTCTGGATGGTGACATATTCGTCA 331
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 249 TGGCCTTAACAGAAATCTCGATCTCTCCCTGGGACAGAGAGGAAATATCCCGCGAT 308
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 332 AGGTCTCAATAAAATCTTGGCTTTAGTCTTGAAGACAGAGAGAGAAATGTTGCGACGAT 391
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 309 TGCTGAGGTGGCTAAGCTTGTGTGTGCTGGTCTGTCTGCTATACCAAGCTTATTTTC 368
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 392 CGCAGAGTGTCTAACTGTTTGCAGATGCTGGCTTAGTGTGCATCACAAGTTTCATATC 451
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 369 TCCATTCGCAAGGATCGTAGAATGCCCCGCAAAATACATGAATCAGCAGGCTGCCATT 428
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 452 ACCCTACACTCAGGATCGCAAAATGCAAGCAAAATTCATGAAGGTGCAAGTTCACCGTT 511
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 429 CTTTGAATATTTGTAGATGACCTCTAAATATTTGTGCAAGCAGAGACGTTAAAGGCT 488
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 512 TTTTGAAGATTTGTGTATGCTCTCTGTCATGTTTGTGAACAGAGGGATGTCGAAGGACT 571
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 489 CTATAAAGGCGCAGAGCTGGGAGATTTAAAGGATTTACAGTATTGATCTGATATGA 548
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 572 CTACAAAAGACCGCGCAGAGAAATTAAGGTTTCACTGGATCGATTTCTGATATGA 631
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 549 GAACCTGAAACTCCTGAGCGTGTCTTTAAACCAATTTGCCAGTGAAGTACTGTGT 608
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 632 AAAGCAGAGCGCCCTGAGTTGGTGTCTGAAACAGACTCTCTGTGATGATAATGACTGTGT 691
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 609 CCACAGGTAGTGAACCTTGCAGAGCAGCAACATTTACCCCTATACATAAACAAGA 668
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 692 CCAGAGTTGTGAACCTTACAGGAACGGGATTTAGTCTGTGGATGCATCTTATGA 751
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 669 TATCAGCAACTTTTGTGCGGAAACAAACTTGCACCAGCTCCGAGCTGAGGCTGAAAC 728
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 752 AGTAAAGAACTATATGTGCCAGAAATAAATCTTATTTGCAAAACAGATGCGGAAC 811
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 729 TCTCCCTTCATTAATTAAGCTGATTAAGCTGATCTCCAGTGGTCCAGGTTTTCGCGAAG 788
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 812 ATTACAGCACTGAAAATTAATAAGTGGATATGCAGTGGTGCAGGTTTTCGCGAAG 871
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 789 CTGGGCCACTCCCTCAAGGTTTCATGCGGAGAGGAGTACTTACAGTTATGCACTT 848
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 872 TTGGCAACCCATGAATGGCTTTATGAGAGAGGAGTACTTGCAGTCCCTTCATTT 931
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 849 TGACACCTGTAGATGTCGCTGATCAACATGAGCATCCCATTTGACTGCGCGTCTC 908
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 932 TGATTTGCTCTGGATGGAGTGTCAATTAATTTGTCTGCTGCTGCTGCTGCTGCTGCT 991
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 909 TGCAGAGATTAAGACAGGCTGGAAGGTGAGCAAGTTTGTCTGCGACATGTTGGACG 968
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 992 TCATGAAGATTAAGAGGCTGGAGCGGTGTACAGCATTTGCTCTGATGTATGAGGCGG 1051
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 969 GAGGTAGTATCTTACGAGCGTGAATTTCTATGACACAGAAAGAGGACCTGTTC 1028
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1052 CCGTGTGGCCATCTTGCATCCAGAGTTTTTTGGACACAGGAAGAGGAGGCTGTGC 1111
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1029 CCGTGTGTGGGGACAACATGTACAAAACACCCCATATCAAAATGGTGTGGAAGTGG 1088
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1112 CAGACAGTGGGAACGACATCCAGAACCCCTTATTAAGATGCTGATGGAACAAGG 1171
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 1089 GGACTGGCTGTTGTTGGAGACCTTCAGTGTCTGGAGAAAATAAGATGAATGATGGCT 1148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1172 AGATTGGCTGATTGGAGGAGATCTTCAAGTCTTGGATCGAGTTTATGGAATGATGGCT 1231
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1149 GGACCAATACCGTCTGACACCTCTCGAGCTCAACAGAAATTAAGAAATGAATGCTGA 1208
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1232 TGATCAGTATGCTCTTACTCTCTAGCTAAAGCAGAAATTTAAAGATATGAATGCTGA 1291
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1209 TGGGTGTTTGCATTCACAGTTGCGCAATCTCTGCTCCACATGCCATGCCCTGTTGATGA 1268
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1292 TGCTGTCTTTCATTCAGTACAGTACGCAACCCAGTGCATGACATGCCCTGTTATGCA 1351
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1269 GGACACCTGCCCGCAGGCTCTAGAGAGGGGTACAAAGCACCCGGTCTCTCTACTACACC 1328
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1352 GGATACCCATTAAGCAACTTCTAGAGAGGGCTACCGCGCCCTGTCCTCTCTCCACCC 1411
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1329 TCTGGCGGCTGGACCAAGGATGAGGATGCGCTTAGCTCTAGACTGGCGGATGAAGCAGCAG 1388
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1412 TCTGGGTGGCTGGACAAAGATGAGGATGTTCTTTGATGTGGCGTATGAAGCAGCATGC 1471
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1389 GGCTGTGCTCGAGGAAGGGTCTCGATGCCCAAGTCAACCATTTGTGCCATCTTTCGTC 1448
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1472 TGCAGTGTGGAGGAGGATGTTGAATCTCTGAGACAGAGTGTGGGCCATCTTCCCATC 1531
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1449 TCCATGTTATATGCTGCCCCACAGAGTCCAGTGGCAGCTCCCGATGATGTC 1508
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1532 TCCCATGATGATGCTGGACCAACTGAGTCCAGTGGCATTCGACAGACGAGATGCTGC 1591
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1509 GGTGGCCAAATTTACATTTGGGAGGAGCCCTCGCAGGAATGCCCATCTCTGAAACCA 1568
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1592 AGGAGCCAACTTTTACATTTTGGACGAGACCCCTGCTGGCATGCTCATCCAGAAACAGG 1651
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1569 GAAGATCTGTATGAACCCACTCATGGGGCAAGTCTTGAGCATGGCCCTGGCTGCAC 1628
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1652 GAAGATCTTATGAGCCAGTATGTTGCCAAAGTCTGAGATGGCCCTGGTGTAT 1711
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1629 CTCTGTGAAATATTCATTCCAGTGGCTGCCCTACACAAAGCAAAAAA 1680
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1712 CACTTTGGAATAGTTCCCTTTTCGAGTTGCAGCTTACACAAAAA 1763
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
RESULT 3
AL540583 951 bp mRNA linear EST 16-FEB-2001
LOCUS AL540583 LTI_FL002_PL1 Homo sapiens cDNA clone CS0DE002YM23 5 prime
DEFINITION , mRNA sequence.
ACCESSION AL540583
VERSION AL540583.1 GI:12870869
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 951)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
FEATURES
source
1. 951
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DE002YM23"
/lab_host="LTI_FL002_PL1"
/clone_lib="LTI_FL002_PL1"
/clone_host="DH10B"
/note="Organ: placenta; Vector: pCMVSPORT 6; 1st strand
cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-stranded cDNA was digested with Not I
and cloned into the Not I and Eco RV sites of the
pCMVSPORT 6 vector. Library was constructed by Life
```


Db 445 CTGCCATCTTTGAATATTTTGTAGATGCACCTCTAAATATTTGTGAAAGCAGACGTA 504
QY 481 AAAGGCTCTATAAAAGGCCAGAGCTGGGAGATTAAAGGATTACAGGTTATGATTC 540
Db 505 AAAGGCTCTATAAAAGGCCAGAGCTGGGAGATTAAAGGATTACAGGTTATGATTC 564
QY 541 GATTATGAGAAACCTGAAACTCCTGAGCGTGTGCTTTAAACCAATTTGTCACAGTGAGT 600
Db 565 GATTATGAGAAACCTGAAACTCCTGAGCGTGTGCTTTAAACCAATTTGTCACAGTGAGT 624
QY 601 GACTGTGTCACAGTAGTGAACCTTCTGCAAGAGCAGAACATTTACCCCTATATATA 660
Db 625 GACTGTGTCACAGTAGTGAACCTTCTGCAAGAGCAGAACATTTACCCCTATATATA 684
QY 661 ATCAAGATATACACGAACCTTTGTGCGGAAACAAACTTGACCACTGCGAGCTGAG 720
Db 685 ATCAAGATATACACGAACCTTTGTGCGGAAACAAACTTGACCACTGCGAGCTGAG 744
QY 721 GCTGAAACTCTCCCTTCATTATCAATTAAGCTGGATCTCCAGTGGTCCAGGTTTG 780
Db 745 GCTGAAACTCTCCCTTCATTATCAATTAAGCTGGATCTCCAGTGGTCCAGGTTTG 804
QY 781 ACAGGAGCTGGGCCACTCCCTCAAAGTTTCATGCGGAGAGAGTAGTACAGTT 840
Db 805 ACAGGAGCTGGGCCACTCCCTCAAAGTTTCATGCGGAGAGAGTAGTACAGTT 864
QY 841 ATGCACCTTTGACACC 855
Db 865 ATGCACCTTTGACACC 879

RESULT 5

BO924373 915 bp mRNA linear EST 20-AUG-2002
LOCUS AGENCOURT_8794733 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6374672
DEFINITION 5', mRNA sequence.

BO924373

BO924373.1 GI:22339404

EST.

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo;
NIH-MGC <http://imgc.ncbi.nih.gov/>;
1 (bases 1 to 915)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: DCTD/DTP/Gazdar

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLCM2554 row: a column: 09

High quality sequence stop: 630.

Location/Qualifiers

FEATURES

Source

1..915

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6374672"

/clone_lib="NIH_MGC_18"

/tissue_type="large cell carcinoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pOT7; Site:1: XhoI; Site:2:
ECORI; cDNA made by oligo-dT priming. Directionally cloned
into EORI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and

Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 252 a 210 c 245 g 207 t 1 others
ORIGIN

Query Match 46.3%; Score 853.6; DB 14; Length 915;
Best Local Similarity 98.1%; Pred. No. 1.2e-239;
Matches 895; Conservative 0; Mismatches 14; Indels 3; Caps 3;

QY 467 AAAGCAGACGCTAAAGGCCCTCTATAAAAGGCCAGAGCTGGGGAGATTAAAGGATTTA 526
Db 1 AAAGCAGACGCTAAAGGCCCTCTATAAAAGGCCAGAGCTGGGGAGATTAAAGGATTTA 60
QY 527 CAGGTATTGATTCTGATTATGAGAAACCTGAAACTCCTGAGCGTGTGCTTTAAACCAATT 586
Db 61 CAGGTATTGATTCTGATTATGAGAAACCTGAAACTCCTGAGCGTGTGCTTTAAACCAATT 120
QY 587 TGTCCACAGTAGTACTGTGTCCACACAGTAGTGAACCTTCTGCAAGAGCAGAACATTG 646
Db 121 TGTCCACAGTAGTACTGTGTCCACACAGTAGTGAACCTTCTGCAAGAGCAGAACATTG 180
QY 647 TACCCTATACTATAATCAAGATATCCACGAACCTCTTTGTGCCGGAACAAACTTGACC 706
Db 181 TACCCTATACTATAATCAAGATATCCACGAACCTCTTTGTGCCGGAACAAACTTGACC 240
QY 707 ACGTCCGAGCTGAGGCTGAAACTCTCCCTTCATTATCAATTACTAAGCTGGATCTCCAGT 766
Db 241 ACGTCCGAGCTGAGGCTGAAACTCTCCCTTCATTATCAATTACTAAGCTGGATCTCCAGT 300
QY 767 GGTCCAGGTTTGTAGCGAAGGCTGGGCCACTCCCTCAAAGGTTTTCATGCGGGAGAGG 826
Db 301 GGTCCAGGTTTGTAGCGAAGGCTGGGCCACTCCCTCAAAGGTTTTCATGCGGGAGAGG 360
QY 827 AGTACTTACAGGTTATGCACTTTTGACACCTCTGTAGATGATGCGCTCAACATCAGCA 886
Db 361 AGTACTTACAGGTTATGCACTTTTGACACCTCTGTAGATGATGCGCTCAACATCAGCA 420
QY 887 TCCCCATTGCTACTGCCCGTCTCTGACAGGATAAGACACGGCTGGAAGGCTGAGCAAGT 946
Db 421 TCCCCATTGCTACTGCCCGTCTCTGACAGGATAAGACACGGCTGGAAGGCTGAGCAAGT 480
QY 947 TGTCTCTGACATGCTGACGGAGGTAGCTATCTTACGAGACGCTGAATTTCTATGAAC 1006
Db 481 TGTCTCTGACATGCTGACGGAGGTAGCTATCTTACGAGACGCTGAATTTCTATGAAC 540
QY 1007 ACAGAAAGAGGAACGCTGTTCCCGTGTGTTGGGGACACATCTACAAAACACCCCATTA 1066
Db 541 ACAGAAAGAGGAACGCTGTTCCCGTGTGTTGGGGACACATCTACAAAACACCCCATTA 600
QY 1067 TCAAAATGCTGATGGAAGTGGGACTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAGA 1126
Db 601 TCAAAATGCTGATGGAAGTGGGACTGGCTGGTGGTGGAGACCTTCAGGTGCTGGAGA 660
QY 1127 AATAGATGGAATGAT-GGGCTGGACCAATACCTCTGACACCTCTGAGAGCTCAACAG 1185
Db 661 AATAGATGGAATGATGGGCTGGACCAATACCTCTGACACCTCTGAGAGCTCAACAG 720
QY 1186 AATGTAAAGAAATGAATGCTGATGCGGTGTTTGGCAATTCAGTTCGCAATCTGTCCAC 1245
Db 721 AATGTAAAGAAATGAATGCTGATGCGGTGTTTGGCAATTCAGTTCGCAATCTGTCCAC 780
QY 1246 AATGGCCATGCCCTGTTGATGAGGACACCTGCCCGAGGCTCTCTAGAGAGGGGCTACAAG 1305
Db 781 AATGGCCCTGCCCTGTTGATGAGGACACCTGCCCGAGGCTCTCTAGAGAGGGGCTACAAG 840
QY 1306 CACCCGGT-CCTCTCTACTACACCTCTGCGGGCTGG-ACCAAGATGAGGATGTCCTC 1363
Db 841 CACCCGGTCTCTCTACTACACCTCTGCGGGCTGGNACCAAGAGGATGACCATGTGCCCTC 900
QY 1364 TAGACTGGCGGA 1375
Db 901 TAGAATGGCCGA 912

361	ATGTGACCTATACATAAAGATATCCAGAACTCTTTGTGCGCGGAAACAACAT	420
703	GACCAGCTCCGAGCTGAGGCTGAAACTCTCCCTTCATTATCAATTAAGCTGGATCTC	762
421	GACCAGCTCCGAGCTGAGGCTGAAACTCTCCCTTCATTATCAATTAAGCTGGATCTC	480
763	CAGTGGGTCCAGGTTTGTAGCGAAGGCTGGGCACTCCCTCAAGGTTTTCATCGCGGAG	822
481	CAGTGGGTCCAGGTTTGTAGCGAAGGCTGGGCACTCCCTCAAGGTTTTCATCGCGGAG	540
823	AAGGAGTACTTACAGGTTATGACATTTGACACCTCTAGATGATGCGCTGATCAACATG	882
541	AAGGAGTACTTACAGGTTATGACATTTGACACCTCTAGATGATGCGCTGATCAACATG	600
883	AGCATCCCATTTGTTACTGCCCTCTCTGCGAGGATGAAGACAGC-GCTGGAAGGTTGCAG	941
601	AGCATCCCATTTGTTACTGCCCTCTCTGCGAGGATGAAGACAGCCTGGAAGGTTGCAG	660
942	CAAGTTTGTCTGGCACATGTTGGAGGAGGTTAGTATCTTACGAGAGCGTGAATTTCTA	1001
661	CAAGTTTGTCTGGCACATGTTGGAGGAGGTTAGTATCTTACGAGAGCGTGAATTTCTA	720
1002	TGAACACAGAAAAGAGGAGCGTGTTCCTGTTTGGGGGACAAACATGTACAAAACACCC	1061
721	TGAACACAGAAAAGAGGAGCGTGTTCCTGTTTGGGGGACAAACATGTACAAAACACCC	780
1062	CCATATCAAAATGGTGATGGAAAGTGGGAGCTGGCTGGTGG 1103	
781	CCATATCAAAATGGTGATGGAAAGTGGGAGCTGGGCTGG 822	
RESULT 7		
BI831135		
LOCUS	603080790F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5172430 5',	
DEFINITION	mRNA sequence.	
ACCESSION	BI831135	
VERSION	BI831135.1 GI:15942685	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	NIH-MGC http://mgc.nci.nih.gov/.	
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)	
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM1429 row: c column: 23 High quality sequence stop: 851.	
FEATURES		
source	Location/Qualifiers	
	1. .860	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:5172430"	
	/tissue_lib="NIH_MGC_119"	
	/tissue_type="medulla"	
	/lab_host="DH10B"	
	/note="Organ: brain; Vector: pCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber	

Db 361 AGTACTACAGGTTATGCACCTTTGACACCCCTGCTAGATGATGCGGTGATCAACATGAGCA 420
QY TCCCCATTGTACTGCGCGTCTCTGCAGAGGATAAGACACGCTGGAGGGTGCAGCAAGT 946
Db TTTTCTCTGCGACATGTTGACGCGTCTCTGCAGAGGATAAGACACGCTGGAGGGTGCAGCAAGT 480
QY TTGTCTCTGCGACATGTTGACGCGTCTCTGCAGAGGATAAGACACGCTGGAGGGTGCAGCAAGT 1006
Db TTTTCTCTGCGACATGTTGACGCGTCTCTGCAGAGGATAAGACACGCTGGAGGGTGCAGCAAGT 540
QY ACAGAAAAGAGGAAACCTGTTCCCGTGTGTTGGGGGACACATGTACAAAACACCCCCATA 1066
Db ACAGAAAAGAGGAAACCTGTTCCCGTGTGTTGGGGGACACATGTACAAAACACCCCCATA 600
QY TCAAAATGTTGATGGAAGTGGGACTGGCTGGTGGTGGAGACCTTCAGGTGCTCTGGAGA 1126
Db TCCAAATGTTGATGGAAGTGGGACTGGCTGGTGGTGGAGACCTTCAGGTGCTCTGGAGA 660
QY AAATAAGATGGAATGATGGGCTGGGACCAATACCGTCTGACACCTCTGGAGCTCAAAACAGA 1186
Db AAATAAGATGGAATGATGGGCTGGGACCAATACCGTCTGACACCTCTGGAGCTCAAAACAGA 720
QY AATGTAAGAAAT-GAATGCTGATGCGGTG-TTTGCATTCAGTTCGCGCAATCC-TGTCC 1243
Db AATGTAAGAAATGGAATGCTGATGCGGTGTTTGTCTTCCAGTTTGGCAATCCTTGTCC 780
QY 1244 ACAATGGCCATGCC 1258
Db 781 CCAATGCCCATGCC 795

RESULT 9
BI829344
LOCUS 603079477F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5171058 5',
DEFINITION mRNA sequence.
ACCESSION BI829344
VERSION BI829344.1 GI:15940894
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1041)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11425 row: j column: 19
High quality sequence stop: 841.
Location/Qualifiers
1..1041
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5171058"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range
0.9-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber

FEATURES
Source

(Invitrogen). Research Genetics tracking code 013. Note:
this is a NIH_MGC Library."
BASE COUNT 269 a 242 c 284 g 246 t
ORIGIN
Query Match 39.8%; Score 734.8; DB 13; Length 1041;
Best Local Similarity 93.1%; Pred. No. 1.2e-204;
Matches 858; Conservative 0; Mismatches 52; Indels 12; Gaps 8;
QY 1 ATCTCGGGGATCAAGAGCAAAAGACCGAGAACCCAGAGAAATCCACCAATGTAGTCTAT 60
Db 57 ATCTCGGGGATCAAGAGCAAAAGACCGAGAACCCAGAGAAATCCACCAATGTAGTCTAT 116
QY 61 CAGCGCCACCATGTGAGCAGGAATAAGAGAGGCAAGTGTGGAACAAGGGTGGGTTTC 120
Db 117 CAGCGCCACCATGTGAGCAGGAATAAGAGAGGCAAGTGTGGAACAAGGGTGGGTTTC 176
QY 121 CGAGGATGTACCGTGTGGCTAACAGGTCTCTCTGGTGTCTGGAACCAACAGTAAGTTTT 180
Db 177 CGAGGATGTACCGTGTGGCTAACAGGTCTCTCTGGTGTCTGGAACCAACAGTAAGTTTT 236
QY 181 GCCTCGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTCTTACTCCCTGGATGGGACAAT 240
Db 237 GCCTCGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTCTTACTCCCTGGATGGGACAAT 296
QY 241 GTCCTCATGGCCTTAACAGAAATCTCGGATTCTCTCTGGGACAGAGAGGAAATATC 300
Db 297 GTCCTCATGGCCTTAACAGAAATCTCGGATTCTCTCTGGGACAGAGAGGAAATATC 356
QY 301 CGCCGGATTGCTGAGGTGGCTAAGCTGTGTCTGTGATGCTGTCTGTCTGCATTACAGC. 360
Db 357 CGCCGGATTGCTGAGGTGGCTAAGCTGTGTCTGTGATGCTGTCTGTCTGCATTACAGC 416
QY 361 TTTATTCTCCATTGCAAGAGATCGTGAGAAATGCCCGCAAAATACATGAATCAGCAGGG 420
Db 417 TTTATTCTCCATTGCAAGAGATCGTGAGAAATGCCCGCAAAATACATGAATCAGCAGGG 476
QY 421 CTGCCATTCTTTGAAA-TATTTGTAGATGCACCTCTAAATATATTTGTGAAAGCAGAGACGT 479
Db 477 CTGCCATTCTTTGAAA-TATTTGTAGATGCACCTCTAAATATATTTGTGAAAGCAGAGACGT 536
QY 480 AAAGGCTCTATAAAGGCCAGAGCTGGGAGATTAAGAGATTTACAGGTATTGATTC 539
Db 537 AAAGGCTCTATAAAGGCCAGAGCTGGGAGATTAAGAGATTTACAGGTATTGATTC 596
QY 540 TGATTATGAAACCTGAAACTCCTGAGCGTGTGCTTAAACCAATTTGTCCACAGTGAG 599
Db 597 TGATTATGAAACCTGAAACTCCTGAGCGTGTGCTTAAACCAATTTGTCCACAGTGAG 656
QY 600 TGACTGTGTCACCGAGTAGTGAACCTC-TGCAAGAGCAGAACATTTACCTATATACTA 658
Db 657 TGACTGTGTCACCGAGTAGTGAACCTCCTTGAAGAGCAGAACATTTACCTATATACTA 716
QY 659 TATC-TAAGATATCAGCAACTTTTGTGCGGAAAACAACTTGACCACTGCCAGCTCCGAGCT 717
Db 717 TAATCAAAAGATATCAGCAACTTTTGTGCGGAAAACAACTTGACCACTGCCAGCTCCGAGCT 776
QY 718 GAGGCTGAAACTCTCCCTTCATTATCAATTT--ACTAAGCTGGATCTCCAGTGGGTCCAGG 775
Db 777 GAGGCTGAAACTCTCCCTTCATTATTTACTAAGCTGGATCTCCAGTGGGTCCAGGT 836
QY 776 TTTTGAGCGAAGCTGGGCCACTC-CCCTCAAGGTTTTCATGCGGGAGAGAGGAGTA---- 830
Db 837 TTTTGAGCGAAGCTGGGCCACTCGCCCTCAAGGGTTTACATGCGGGAGAGAGGAGGAC 896
QY 831 CTATACAGGTATGCACCTTTGACACCTTGCT-AGATGATGGCGTGA-TCAACATGAGCATC 888
Db 897 TTACCAAGGTATGCCATTTGAACCCCTGTTAACAATGAAGGGGTGAGACAATTTGAGCGTG 956
QY 889 CCCATTGTACTGCCCTCTCTGTG 910
Db 957 CCGGTGTGACTGCCGCGCTATG 978

```

RESULT 10
BQ942867      1054 bp      mRNA      linear      EST 21-AUG-2002
LOCUS        AGENCOURT_8836984 NIH_MGC_101 Homo sapiens cDNA clone IMAGE:6429378
DEFINITION   5', mRNA sequence.
ACCESSION    BQ942867
VERSION      BQ942867.1 GI:22358345
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
              1 (bases 1 to 1054)
REFERENCE    NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished (1999)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Rubin Laboratory
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LCM2617 row: h column: 19
              High quality sequence start: 24
              High quality sequence stop: 736.
              Location/Qualifiers
                1..1054
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6429378"
                  /clone_lib="NIH_MGC_101"
                  /tissue_type="epidermoid carcinoma, cell line"
                  /lab_host="DH10B (phage-resistant)"
                  /note="Organ: lung; Vector: pOTB7; Site:1: EcoRI; Site:2:
                  XhoI; cDNA made by oligo-dT priming. Directionally cloned
                  into ECORI/XhoI sites using the following 5' adaptor:
                  GGCAGGAG(G). Library constructed by Ling Hong in the
                  laboratory of Gerald M. Rubin (University of California,
                  Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and a
                  Superscript II RT (Life Technologies). Note: this is a
                  NIH_MGC Library."
BASE COUNT   285 a 231 c 279 g 254 t 5 others
ORIGIN
Query Match 39.2%; Score 722.6; DB 14; Length 1054;
Best Local Similarity 94.4%; Pred. No. 4.7e-201;
Matches 806; Conservative 0; Mismatches 35; Indels 13; Gaps 5;

QY 180 TGCCCTGAGGAGTACCTT-GTCTCCCATGCCATCCCTGTGTACTCCCTGGATGGGACA 238
Db 19 TGGCTGTGAGGAGTACCTGTGCTCCCATGCCATCCCTGTGTACTCCCTGGATGGGACA 78
QY 239 ATGTCGCTCATGGCTTAACGAATCTCGGATCTCTCTCTGGGACAGAGAGAAATA 298
Db 79 ATGTCGCTCATGGCTTAACGAATCTCGGATCTCTCTCTGGGACAGAGAGAAATA 138
QY 299 TCCGCGGATGCTGAGTGCGTAAAGCTGTGCTGATGCTGGCTGTGCTGCAATTACA 358
Db 139 TCCGCGGATGCTGAGTGCGTAAAGCTGTGCTGATGCTGGCTGTGCTGCAATTACA 198
QY 359 GCTTATTCTCCATCCCAAGGATCGTGAATCCCGCAAAATACATCAATCAGCAG 418
Db 199 GCTTATTCTCCATCCCAAGGATCGTGAATCCCGCAAAATACATCAATCAGCAG 258
QY 419 GGCTGCCATCTTTGAAATATTGTAGATGACCTCTAAATATTGTGAAAGCAGACG 478
Db 259 GGCTGCCATCTTTGAAATATTGTAGATGACCTCTAAATATTGTGAAAGCAGACG 318
QY 479 TAAAGGCCCTCTATAAAGGCCAGAGCTGGGAGATTAAAGGATTTACAGGTATTGATT 538

```

```

Db 319 TAAAGGCCCTCTATAAAGGCCAGAGCTGGGAGATTAAAGGATTTACAGGTATTGATT 378
QY 539 CTGATTATGAGAAACCTCAAACTCCTGAGCGTGTCTTAAAAACCAATTTGTCCACAGTGA 598
Db 379 CTGATTATGAGAAACCTCAAACTCCTGAGCGTGTCTTAAAAACCAATTTGTCCACAGTGA 438
QY 599 GTGACTGTGTCCACCAGGTAGTGGAACTTCTCAAGAGCAGAACTTTGACCCCTATACATA 658
Db 439 GTGACTGTGTCCACCAGGTAGTGGAACTTCTCAAGAGCAGAACTTTGACCCCTATACATA 498
QY 659 TAAACAAGATATCCACCAACTCTTTGTCGGGAAACAACTTGCACACGTCGCCAGCTG 718
Db 499 TAAACAAGATATCCACCAACTCTTTGTCGGGAAACAACTTGCACACGTCGCCAGCTG 558
QY 719 AGGCTGAACACTCTCCCTTCATTATCAATTAATCAATTAATCAATTAATCAATTAATCA 778
Db 559 AGGCTGAACACTCTCCCTTCATTATCAATTAATCAATTAATCAATTAATCAATTAATCA 618
QY 779 TGAGCAAGGCTGGCCCACTCCCTCAAGGTTTCATCGGGAGAGAGTACTTACAGG 838
Db 619 TGAGCAAGGCTGGCCCACTCCCTCAAGGTTTCATCGGGAGAGAGTACTTACAGG 678
QY 839 TTATGCACCTTTGACACCTCTGCTAGATGATGGCTGATCAACATGAGCATCCCATTTGAC 898
Db 679 TTATGCACCTTTGACACCTCTGCTAGATGATGGCTGATCAACATGAGCATCCCATTTGAC 738
QY 899 TG-CCTCTCTGTCAGAGGATAAGACAGCGGTGGAAGGTCGAGCAAG-----TTTGTCC 952
Db 739 TGCCCTGCTCTGTCAGAGGATAAGACAGCGGTGGAAGGTCGAGCAAGTTTGTCTCC 798
QY 953 TGGACATGTTGGTGGACGAGGG-----TAGCTATCTTACGAGACGCTGAATTTAT-GAAC 1006
Db 799 TGGACATGTTGGTGGACGAGGGTAACCTATTCTTACGAGGACGCTGAATTTATGGAAC 858
QY 1007 ACAGAAAAGAGGAA 1020.
Db 859 ACAGAAAAGAGGGA 872

RESULT 11
BQ943498
LOCUS        AGENCOURT_8785245 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6370861
DEFINITION   5', mRNA sequence.
ACCESSION    BQ943498
VERSION      BQ943498.1 GI:22358976
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
              1 (bases 1 to 874)
REFERENCE    NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS     National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE       Unpublished (1999)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: DCTD/DP/GenZdar
              cDNA Library Preparation: Rubin Laboratory
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LCM2544 row: b column: 14
              High quality sequence stop: 688.
              Location/Qualifiers
                1..874
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:6370861"
                  /clone_lib="NIH_MGC_18"
FEATURES
source

```

```
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT      210 a   227 c   244 g   193 t
ORIGIN
Query Match      38.2%; Score 704.8; DB 14; Length 874;
Best Local Similarity 96.1%; Pred. No. 7.5e-196;
Matches 761; Conservative 0; Mismatches 12; Indels 19; Gaps 3;
QY 665 AAGATATCCAGCACTCTTTGTCGCGGAAACAACTTGACCACTGCGAGCTGAGCTG 724
DB 1 AAGATATCCAGCACTCTTTGTCGCGGAAACAACTTGACCACTGCGAGCTGAGCTG 60
QY 725 AAATCTCCCTTCAATTAATTAAGCTGGATCTCCAGTGGTCCAGGTTTGAGCG 784
DB 61 AAATCTCCCTTCAATTAATTAAGCTGGATCTCCAGTGGTCCAGGTTTGAGCG 120
QY 785 AAGCTGGGCACTCCCTCAAGGTTTCATGCGGGAGAGGAGTACTTACAGGTTATGC 844
DB 121 AAGCTGGGCACTCCCTCAAGGTTTCATGCGGGAGAGGAGTACTTACAGGTTATGC 180
QY 845 ACTTTGACACCTGCTAGA-----TGATGGCGTGCATCAATGAGCATCC 889
DB 181 ACTTTGACACCTGCTAGATGGATGCGCCCTTCTGATGGCGTGCATCAATGAGCATCC 240
QY 890 CCATTTGACTGCCGCTCTCTGCAGAGATGAAGACACGCGTGGAGGGTGCAGCAAGTTTG 949
DB 241 CCATTTGACTGCCGCTCTCTGCAGAGATGAAGACACGCGTGGAGGGTGCAGCAAGTTTG 300
QY 950 TCCTGGCACATGTTGGAGGAGGAGTACTATCTTACGAGCGCTGAATTCATGAACACA 1009
DB 301 TCCTGGCACATGTTGGAGGAGGAGTACTATCTTACGAGCGCTGAATTCATGAACACA 360
QY 1010 GAAAGAGGAACGCTGTTCCTCCGTGTTTGGGGGACAACTATACAAAACCCCAATCA 1069
DB 361 GAAAGAGGAACGCTGTTCCTCCGTGTTTGGGGGACAACTATACAAAACCCCAATCA 420
QY 1070 AAATGGTATGGAAGTGGGAGCTGGCTGTTGGTGAGACCTTCAGGTGCTGGAGAAA 1129
DB 421 AAATGGTATGGAAGTGGGAGCTGGCTGTTGGTGAGACCTTCAGGTGCTGGAGAAA 480
QY 1130 TAAGATGGAATGATGGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAGAAAT 1189
DB 481 TAAGATGGAATGATGGCTGGACCAATACCGTCTGACACCTCTGGAGCTCAACAGAAAT 540
QY 1190 GTAAGAAATGAATGTGATGCGGTGTTTGCATTCCAGTTGCGCAATCCCTGCCAATG 1249
DB 541 GTAAGAAATGAATGTGATGCGGTGTTTGCATTCCAGTTGCGCAATCCCTGCCAATG 600
QY 1250 GCCATGCCCTGTTGATGCAGGACACCTGCGCGAGGCTCTCTAGAGAGGGGCTACAAGCACC 1309
DB 601 GCCATGCCCTGTTGATGCAGGACACCTGCGCGAGGCTCTCTAGAGAGGGGCTACAAGCACC 660
QY 1310 CGGTCTCTCTACTACACCTCTGGGCGGTGGACCAAGGATGACGATGCGCTCTAGACT 1369
DB 661 CGGTCTCTCTACTACACCTCTGGGCGGTGGACCAAGGATGACGATGCGCTCTAGACT 720
QY 1370 GCGGGATGAAGCAGCAGCG--GGCTGTGCTCGAGCAAGGGGTGC---TGGATCCCAAGTCA 1425
DB 721 GCGGGATGAAGCAGCAGCGGGGCTGTGCTCCGAGAGAGGGGCTCTGGGATCCCAAGTCA 780
QY 1426 ACCATTGTTGCC 1437
DB 781 CCATTGTTGCC 792
```

```
RESULT 12
B0935947
LOCUS      917 bp      mRNA      linear      EST 21-AUG-2002
DEFINITION AGENCOURT_8772851 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6372309
5', mRNA sequence.
ACCESSION B0935947
VERSION B0935947.1 GI:22351330
KEYWORDS EST.
SOURCE      human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 917)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2547 row: n column: 22
High quality sequence stop: 585.
FEATURES
Location/Qualifiers
1..917
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6372309"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT      223 a   225 c   252 g   214 t   3 others
ORIGIN
Query Match      36.4%; Score 672.4; DB 14; Length 917;
Best Local Similarity 96.7%; Pred. No. 2.6e-186;
Matches 707; Conservative 0; Mismatches 19; Indels 5; Gaps 2;
QY 1 ATGTGGGGATCAAGAGCAAAAGACGAGAGAACCCAGAGAACCCACCAATGTAGTCTAT 60
DB 187 ATGTGGGGATCAAGAGCAAAAGACGAGAGAACCCAGAGAACCCACCAATGTAGTCTAT 246
QY 61 CAGGCCCACTGTGACGAGGAATAAGAGAGGCAAGTGGTGGAAACAAGGGGGTTC 120
DB 247 CAGGCCCACTGTGACGAGGAATAAGAGAGGCAAGTGGTGGAAACAAGGGGGTTC 306
QY 121 CGAGGATGTACCTGTGGCTAACAGGTCTCTCTGGTCTCGAAACCAACGATAAGTTT 180
DB 307 CGAGGATGTACCTGTGGCTAACAGGTCTCTCTGGTCTCGAAACCAACGATAAGTTT 366
QY 181 GCCTGTGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAT 240
DB 367 GCCTGTGAGGAGTACCTTGTCTCCCATGCCATCCCTTGTACTCCCTGGATGGGACAAT 426
QY 241 GTCCGTATGGCCTTAAAGAAATCTCGGATTCCTCTGGGACAGAGAGGAAATATC 300
DB 427 GTCCGTATGGCCTTAAAGAAATCTCGGATTCCTCTGGGACAGAGAGGAAATATC 486
QY 301 CGCCGGATTCGTGAGTGGCTAAGCTGTTTGGCTGATGCTGCTGGTCTGCATTACCAAGC 360
```


JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DMP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9639 row: c column: 11
High quality sequence stop: 648.

FEATURES

source
1..650
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3877186"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.8 kb. Library constructed by Life
Technologies."

BASE COUNT 184 a 145 c 170 g 151 t

ORIGIN

Query Match 33.9%; Score 624.8; DB 12; Length 650;
Best Local Similarity 99.4%; Pred. No. 2.3e-172;
Matches 648; Conservative 0; Mismatches 2; Indels 2; Gaps 2;
QY 538 TCTGATTATGAGAACTGAACTCCCTGAGCGTGTGCTTAAACCAATTTGTCCACAGTG 597
Db 1 TCTGATTATGAGAACTGAACTCCCTGAGCGTGTGCTTAAACCAATTTGTCCACAGTG 60
QY 598 AGTGACTGTGTCCACAGGAGTGTGGAAGTCTGTCAAGAGCAGACATGTGACCTATACT 657
Db 61 AGTGACTGTGTCCACAGGAGTGTGGAAGTCTGTCAAGAGCAGACATGTGACCTATACT 120
QY 658 ATAATCAAGATATCCACGACTTTTGTGCGGGAACAACTTGACACGTCCTCGAGCT 717
Db 121 ATAATCAAGATATCCACGACTTTTGTGCGGGAACAACTTGACACGTCCTCGAGCT 180
QY 718 GAGGCTGAACTCTCCCTTCATTATCAATTAAGCTGATCTCCAGTGGTCCAGGTT 777
Db 181 GAGGCTGAACTCTCCCTTCATTATCAATTAAGCTGATCTCCAGTGGTCCAGGTT 240
QY 778 TTGAGCGAAGGCTGGGCCACTCCCTCAAGGTTTTCATGCGGGAAGAGTACTTACAG 837
Db 241 TTGAGCGAAGGCTGGGCCACTCCCTCAAGGTTTTCATGCGGGAAGAGTACTTACAG 300
QY 838 GTTATGCACATTGACACCTCTAGATGATGGCGTGATCAACATGAGCATCCCATTTGA 897
Db 301 GTTATGCACATTGACACCTCTAGATGATGGCGTGATCAACATGAGCATCCCATTTGA 360
QY 898 CTGCGCGTCTCTCAGAGGATAGACACGCTGGAAGGTCGAGCAAGTTTGTCTTGCGCA 957
Db 361 CTGCGCGTCTCTCAGAGGATAGACACGCTGGAAGGTCGAGCAAGTTTGTCTTGCGCA 420
QY 958 CATGGTGGAGGAGGTAGTATCTTACGAGACGCTGAATCTTATGAACACAGAAAGAG 1017
Db 421 CATGGTGGAGGAGGTAGTATCTTACGAGACGCTGAATCTTATGAACACAGAAAGAG 480
QY 1018 GAACGCTGTTCCTGTTTGGGGGACAACATGTACAAAACACCCCATATCAAAATGGTG 1077
Db 481 GAACGCTGTTCCTGTTTGGGGGACAACATGTACAAAACACCCCATATCAAAATGGTG 540
QY 1078 ATGGAAGTGGGACTGGCTGTGGTGGAGACCTTCAGTGTGGAGAAAATAAGATCG 1137
Db 541 ATGGAAGTGGGACTGGCTGTGGTGGAGACCTTCAGTGTGGAGAAAATAAGATCG 600
QY 1138 AATGATGGGCTGGCAATACCGTCTGACACCTCTGGAGCTCAAAACAGAAAT 1189
|||||

Db

601 AATGATGGGCTGGACCAATA-CGTCTGACA-CTCTGGAGCTCAACAGAAAT 650

RESULT 15

BI103516

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

THIS PAGE BLANK (11/27/70)